

## Article

# An Online Iterative Linear Quadratic Approach for a Satisfactory Working Point Attainment at FERMI

Niky Bruchon <sup>1,\*</sup>, Gianfranco Fenu <sup>1</sup>, Giulio Gaio <sup>2</sup>, Simon Hirlander <sup>3</sup>, Marco Lonza <sup>2</sup>, Felice Andrea Pellegrino <sup>1</sup> and Erica Salvato <sup>1</sup>

<sup>1</sup> Department of Engineering and Architecture, University of Trieste, 34127 Trieste, TS, Italy; fenu@units.it (G.F.); fapellegrino@units.it (F.A.P.); erica.salvato@phd.units.it (E.S.)

<sup>2</sup> IT Group, Elettra Sincrotrone Trieste, 34149 Basovizza, TS, Italy; giulio.gαιο@elettra.eu (G.G.); marco.lonza@elettra.eu (M.L.)

<sup>3</sup> IDA Lab Salzburg, University of Salzburg, 5020 Salzburg, Austria; simon.hirlander@sbg.ac.at

\* Correspondence: niky.bruchon@phd.units.it

† Current address: Institute for Beam Physics and Technology, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany.

**Abstract:** The attainment of a satisfactory operating point is one of the main problems in the tuning of particle accelerators. These are extremely complex facilities, characterized by the absence of a model that accurately describes their dynamics, and by an often persistent noise which, along with machine drifts, affects their behaviour in unpredictable ways. In this paper, we propose an online iterative Linear Quadratic Regulator (iLQR) approach to tackle this problem on the FERMI free-electron laser of Elettra Sincrotrone Trieste. It consists of a model identification performed by a neural network trained on data collected from the real facility, followed by the application of the iLQR in a Model-Predictive Control fashion. We perform several experiments, training the neural network with increasing amount of data, in order to understand what level of model accuracy is needed to accomplish the task. We empirically show that the online iLQR results, on average, in fewer steps than a simple gradient ascent (GA), and requires a less accurate neural network to achieve the goal.

**Keywords:** neural networks; optimization; control systems; free-electron laser



**Citation:** Bruchon, N.; Fenu, G.; Gaio, G.; Hirlander, S.; Lonza, M.; Pellegrino, F.A.; Salvato, E. An Online Iterative Linear Quadratic Approach for a Satisfactory Working Point Attainment at FERMI. *Information* **2021**, *12*, 262. <https://doi.org/10.3390/info12070262>

Academic Editor: Gianluca Valentino

Received: 21 May 2021

Accepted: 19 June 2021

Published: 26 June 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

A Free-Electron Laser (FEL) [1] is a highly tunable source of coherent radiation. Indeed, it has the widest wavelength range of all laser types, currently ranging from microwaves to X-rays. Such light sources are used for cutting-edge research in many different fields of physics, chemistry, and biology. The FEL principle is based on a beam of electromagnetic waves traveling collinearly with an electron beam inside a magnetic structure called modulator undulator. As a result of their superimposition, the electron beam is energy modulated. The energy modulation is then converted into a charge-density modulation that produces micro-bunching in the longitudinal direction according to the radiation wavelength. Finally, the micro-bunched electron beam radiates in phase with the electromagnetic waves amplifying the radiation until saturation. With a sufficiently high gain, the amplified radiation continues to interact with the electron beam traveling together with the radiation. This configuration is called Self-Amplified Spontaneous Emission (SASE) [2] and produces partially coherent radiation starting from an incoherent spontaneous emission. Even if the radiation has high peak power and good transverse coherence at saturation, the longitudinal coherence is affected by strong pulse-to-pulse spectro-temporal fluctuations. An improvement of the longitudinal coherence is provided by seeded FELs [3–6] where an external optical laser is injected with the electron beam in the modulator undulator. In this case, the energy modulation has better properties since it is seeded by the optical laser and

not by incoherent spontaneous emission. Furthermore, a shorter magnetic path is required to amplify the output until saturation. As a result, a crucial parameter in seeded FELs is the longitudinal and transverse overlapping between electrons and laser, which is obtained by setting a proper working point.

In general, FEL optimization has always been a very demanding task. The system complexity, the lack of a complete theoretical model, the noise, and the unpredictability of machine drifts make it even harder. Over the years, several approaches have been proposed, some of them being just as a proof of principle, others being actually employed [7].

A model-free approach using Gradient Ascent and Extremum Seeking algorithms has been investigated on the FERMI FEL at Elettra Sincrotrone Trieste [8]. Furthermore, a multi-physics simulation tool kit called OCELOT [9] has been designed at the European XFEL in Hamburg for the study of FELs and synchrotron light sources. Some generic optimization algorithms such as Extremum Seeking, Nelder Mead, and Bayesian optimization based on Gaussian processes are already implemented in the framework. Gaussian process and Bayesian optimization have been also employed to tune the quadrupole currents at the Stanford Linear Accelerator Center (SLAC) [10,11] and to optimize the self-amplification power of the spontaneous emission of the Free electron LASer in Hamburg (FLASH) at the Deutsches Elektronen-Synchrotron (DESY) [12]. In recent years, Machine Learning (ML) techniques have led to many improvements and successful implementations in the field of particle accelerators, from automated alignment of various devices with beam to optimising different parameters, see for example [13–18]. An overview of the opportunities provided by the application of ML for particle physics is given in [19]. In particular, the authors of [20,21] consider Reinforcement Learning (RL) for control and performance improvement. In [22], the authors advocate the use of artificial neural networks to model and control particle accelerators in combination with RL. Moreover, recent works [23–29] have presented RL methods used in the context of FELs. In [23,24], the FEL model and the policy are defined by neural networks in a simulation environment, while in [25] the authors present an application of RL on a real system.

In the present paper, we face the problem of the attainment of a satisfactory working point, i.e., the one resulting in a detected intensity greater than a given threshold. In particular, we are concerned with the alignment of the seed laser to properly tune the superimposition with the electron beam in the modulator undulator of the FERMI FEL. Such a challenging task has already been approached by some of the authors employing RL, both in its model-free and model-based variants [26,27,29]. Here, we investigate the use of an iterative Linear Quadratic Regulator (iLQR) approach [30] to deal with the problem at hand. That algorithm has been already successfully employed in the control of non-linear biological movement systems [30], and recently proposed for the trajectory correction on the Advanced Proton Driven Plasma Wakefield Acceleration Experiment (AWAKE) at CERN [31]. We describe the FEL dynamics using a state-space representation [32], where the association between the position of two tip-tilts and the resulting intensity is identified by training a Neural Network (NN) on real data samples [33–35]. We propose to use the iLQR in a Model-Predictive Control (MPC) fashion in order to deal with possible model-plant mismatches. We refer to the mentioned scheme as online iLQR. We evaluate the effectiveness of such an approach as the amount of data used for the NN training increases. Moreover, we compare its performances with a gradient ascent (GA), showing that the online iLQR needs fewer control steps for task achievement.

Clearly, training an NN is not the only way to identify a model of a nonlinear function, such as a regressor. Other possible approaches, for example, rely on Gaussian–Laplace [36] or exponential [37] models. However, they are limited to those problems in which it is possible to assume that the function to be approximated has a known shape.

The paper is organized in the following way. Section 2 presents a general description of the seed laser alignment system, as well as the proposed algorithm. In Section 3, the experimental configuration and the achieved results are reported and discussed. Finally, conclusions are drawn in the last section.

## 2. Materials and Methods

In the following, we first describe the main elements of FERMI involved in the considered task (Section 2.1). Subsequently, we recall the iterative Linear Quadratic Regulator (iLQR) method (Section 2.2) and describe an online implementation (Section 2.3), employed for the locally optimal feedback control of the FERMI process.

### 2.1. FERMI Facility

Despite the complex structure of the FERMI, the alignment process can be described by the simple setup shown in Figure 1. Assuming a stable electron beam trajectory, the superimposition of the two beams is achieved by imposing a specific laser trajectory. In particular, by properly tuning the angle of two tip-tilt mirrors upstream of the facility (TT1 and TT2), a coherent optical radiation downstream of the chain is detectable by the intensity sensor ( $I_0$  monitor). For a detailed description of the alignment process, and of each of the devices shown in Figure 1, we refer readers to [26,27]. Here, we only point-out that the intensity detected by the  $I_0$  monitor can be adjusted by properly controlling the pitch and yaw movements of the tip-tilts, which depend on the voltage regulation of some piezomotors (two for each tip-tilt mirror).

We denote by  $v_{(k)}^{\text{pitch},i}$  and  $v_{(k)}^{\text{yaw},i}$  the servo motors voltages at the  $k$ -th time instant, governing, respectively, the pitch and yaw angles of the  $i$ -th tip-tilt mirror. The angles are controlled incrementally, i.e., at each  $k$ -th time instant, the  $i$ -th tip-tilt mirror receives the pitch and yaw displacements as input (respectively, denoted by  $\delta v_{(k)}^{\text{pitch},i}$  and  $\delta v_{(k)}^{\text{yaw},i}$ ), and reaches the new  $v_{(k+1)}^{\text{pitch},i}$  and  $v_{(k+1)}^{\text{yaw},i}$  according to:

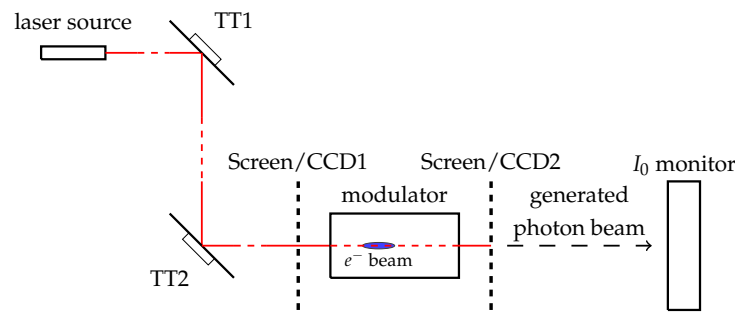
$$\begin{aligned} v_{(k+1)}^{\text{pitch},i} &= v_{(k)}^{\text{pitch},i} + \delta v_{(k)}^{\text{pitch},i} \\ v_{(k+1)}^{\text{yaw},i} &= v_{(k)}^{\text{yaw},i} + \delta v_{(k)}^{\text{yaw},i} \end{aligned} \tag{1}$$

Let  $X$ ,  $U$  and  $I$  be, respectively, the state set, the control set and the output set of the system represented in Figure 1. We model it as a discrete-time dynamical system  $\mathcal{S}$  whose dynamics follows:

$$\begin{cases} x_{(k+1)} &= Ax_{(k)} + Bu_{(k)} \\ I_{(k)} &= f(x_{(k)}) \end{cases}, \tag{2}$$

where  $x_{(k)} := [v_{(k)}^{\text{pitch},1} \ v_{(k)}^{\text{yaw},1} \ v_{(k)}^{\text{pitch},2} \ v_{(k)}^{\text{yaw},2}]^\top \in X \subset \mathbb{R}^4$  is the state at the  $k$ -th time instant,  $u_{(k)} := [\delta v_{(k)}^{\text{pitch},1} \ \delta v_{(k)}^{\text{yaw},1} \ \delta v_{(k)}^{\text{pitch},2} \ \delta v_{(k)}^{\text{yaw},2}]^\top \in U \subset \mathbb{R}^4$  is the control input at the  $k$ -th time instant,  $I_{(k)} \in I \subset \mathbb{R}$  is the detected intensity at the  $k$ -th time instant,  $A, B \in \mathbb{R}^4 \times \mathbb{R}^4$  are two identity matrices, and  $f : X \rightarrow I$  is a non linear function. In other words, it is a Wiener system [38], i.e., a model consisting of a static nonlinear element preceded by a linear dynamic system.

For our purpose, we represent  $f$  through an NN trained on data collected at the beginning of the whole procedure, as described in Section 3. Clearly, the dynamical system  $\mathcal{S}$  (Equation (2)) is only a simplified approximation of the FERMI dynamics. Therefore, the employed control strategy must be take into account possible model-plant mismatches.



**Figure 1.** Simple scheme of the FERMI FEL seed laser alignment set up. TT1 and TT2 are the tip–tilt mirrors, Screen/CCD1 and Screen/CCD2 are two removable Yttrium Aluminum Garnet (YAG) screens with Charge Coupled Devices (CCDs) and  $I_0$  monitor is the employed intensity sensor.

### 2.2. Iterative Linear Quadratic Regulator

Consider a discrete-time dynamical system of the form:

$$x_{(k+1)} = g\left(x_{(k)}, u_{(k)}\right), \tag{3}$$

where  $x_{(k)} \in X$  and  $u_{(k)} \in U$  are, respectively, the state and the control input at the  $k$ -th time instant, while  $g : X \times U \rightarrow X$  is a non linear function.

The Iterative Linear Quadratic Regulator [30] is an iterative procedure used to employ the Linear Quadratic Regulator (LQR) approach [39] to a non-linear system (Equation (3)). It results in a locally-optimal state-feedback control law able to minimize, in a finite horizon  $H$ , a quadratic cost of the form:

$$J = \frac{1}{2} \sum_{k=0}^{H-1} \left[ \left(x_{(k)} - x^*\right)^\top Q \left(x_{(k)} - x^*\right) + u_{(k)}^\top R u_{(k)} \right] + \frac{1}{2} \left(x_{(H)} - x^*\right)^\top Q_H \left(x_{(H)} - x^*\right), \tag{4}$$

where  $Q = Q^\top \succeq 0$ ,  $Q_H = Q_H^\top \succeq 0$  and  $R = R^\top \succ 0$  are, respectively, the state cost matrix, the final state cost matrix, and the input cost matrix, while  $x^*$  is the desired terminal state (we label positive (semi-)definiteness  $\succ (\succeq)$ ).

Let  $\mathbf{x}_i = \left(x_{(0)}, x_{(1)}, \dots, x_{(H)}\right)$  be an  $H$  step state trajectory, result of a control sequence  $\mathbf{u}_i = \left(u_{(0)}, u_{(1)}, \dots, u_{(H-1)}\right)$  at the  $i$ -th iteration of the procedure. The corresponding cost-to-go value  $J_i$  is obtained according to Equation (4). At each  $i$ -th iteration of the procedure:

1. the non-linear system is linearized around a nominal trajectory  $\mathbf{x}_{i-1}$ , and the result of a nominal control sequence  $\mathbf{u}_{i-1}$  applied to the open-loop system;
2. a finite-horizon LQR problem is solved, providing as output a new control sequence  $\text{vec}(\mathbf{u}_i) = \text{diag}\left\{K_j^{\text{LQR}}, j = 0 \dots H\right\} \text{vec}(\mathbf{x}_i)$ ; where  $K_j^{\text{LQR}}$  are the gain matrices that solve the finite-horizon LQR problem and  $\text{vec}(\cdot)$  is the vector operator.

The two steps are repeated, starting from the new trajectory  $\mathbf{x}_i$ , until convergence, i.e., until the difference of the two subsequent costs ( $J_i - J_{i-1}$ ) is lower than a tolerance (tol). The resulting output is an optimal control sequence  $\mathbf{u}^* = \left(u_{(0)}^*, u_{(1)}^*, \dots, u_{(H-1)}^*\right)$ . More details of the iterative procedure can be found in [30].

### 2.3. Online iLQR and Augmented State

In order to apply the iLQR approach to the problem at hand, two adjustments are needed.

First, the iLQR procedure described above is essentially open-loop, and thus it is potentially affected by modeling errors and disturbances. In order to provide the iLQR with the ability to properly address possible dynamical deviations between Equation (3) and the real plant to be controlled, we employ iLQR in an MPC fashion [40]. In other words, only the first input of the optimal control sequence  $\mathbf{u}^*$  is actually applied, and the whole op-

timization is repeated at the subsequent time step. At each  $k$ -th time instant, the iLQR is fed by the state of the system  $x_{(k)}$  and outputs a control sequence  $\mathbf{u}^* = (u_{(k)}^*, \dots, u_{(k+H)}^*)$ , the result of the iterative procedure performed on the internal dynamical model Equation (3) (where  $H$  is the LQR horizon). Only the first control input  $u_{(k)}^*$  is applied to the system, and the process is repeated from the new state.

Second, to allow the cost function in Equation (4) to depend on the intensity, we employ an augmented state  $\hat{x}_{(k)} = [x_{(k)}^\top \ I_{(k)}]^\top$ , thus the model state equation becomes:

$$\begin{bmatrix} x_{(k+1)} \\ I_{(k+1)} \end{bmatrix} = \begin{bmatrix} Ax_{(k)} + Bu_{(k)} \\ f(x_{(k)}) + h(x_{(k)}, u_{(k)}) \end{bmatrix}, \tag{5}$$

where  $h(x_{(k)}, u_{(k)})$  represents the change of intensity due to the input  $u_{(k)}$  occurring at time  $k$  when the tip-tilt mirrors are in  $x_{(k)}$ . From the practical standpoint, when solving the iLQR, the function  $h$  is computed as:

$$h(x_{(k)}, u_{(k)}) = \nabla f [x_{(k)}^\top \ u_{(k)}^\top]^\top, \tag{6}$$

where  $\nabla$  denotes the gradient.

### 3. Implementations and Results

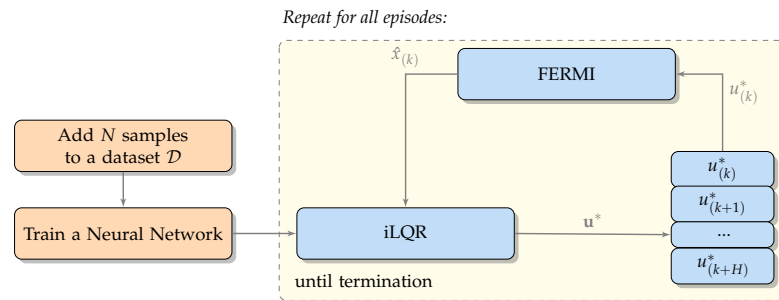
We next describe the conducted experiments (Section 3.1), and then show the obtained results (Section 3.2). In particular, we investigate the behavior of the iLQR as the amount of data provided to the NN during training increases. Indeed, the iLQR relies on the model Equation (2), whose (static) nonlinearity is described by an NN. Thus, the first step of the procedure is, essentially, an identification step, in which input–output pairs are collected and employed to train the network, i.e., to approximate the function  $f$  of Equation (2). Moreover, we compare the proposed approach to a simple GA approach, which employs the same approximation.

#### 3.1. Experimental Procedure

We perform the online iLQR procedure of Section 2.3 on the real FERMI facility. We model the alignment dynamics at FERMI with  $\mathcal{S}$  (Equation (2)), where  $f$  is identified as explained in the following, and use it for performing each iLQR step of the online procedure. The resulting control inputs will be denoted by  $u_{(k)}^* := [\delta v_{(k)}^{\text{pitch},1}, \delta v_{(k)}^{\text{yaw},1}, \delta v_{(k)}^{\text{pitch},2}, \delta v_{(k)}^{\text{yaw},2}]^\top$ . We set a maximum number of steps  $N_{\text{max}}$  to obtain at least the 95% of a target intensity  $I^*$  on the  $I_0$  monitor. Hereafter, we will use  $x_{(k)}^R$  and  $I_{(k)}^R$  to denote the actual state and the actual detected intensity at the  $k$ -th time instant, while  $x_{(k)}$  and  $I_{(k)}$  are their respective counterparts in  $\mathcal{S}$ . The pitch and yaw displacement performed at the  $k$ -th time instant on the FERMI are denoted by  $u_{(k)}^R$ . The experimental procedure consists of 6 subsequent runs of 30 episodes. At the beginning, a set of 30 initial states is randomly selected, to be used as initial conditions of the 30 episodes for all the 6 runs. A target intensity  $I^*$  is set according to the operating condition of the manually tuned plant. The associated tip-tilts' state is denoted by  $x^*$ .

Each run starts collecting  $N$  random  $(x_{(i)}^R, I_{(i)}^R)$ -samples, i.e., state and intensity pairs: we initialize the FERMI in a random initial state  $x_{(0)}^R$ , resulting in a detected  $I_{(0)}^R$  on the  $I_0$  monitor, and we collect the data-set  $\mathcal{D} = \{(x_{(0)}^R, I_{(0)}^R), \dots, (x_{(N)}^R, I_{(N)}^R)\}$  by performing random pitch and yaw displacements  $\{u_{(0)}^R, \dots, u_{(N-1)}^R\}$  on the two tip-tilts' mirrors. The data collected at the beginning of the current run, are added to the data collected so far, in such a way that the dataset becomes richer as new runs are performed. The collected data-set  $\mathcal{D}$  is then used to train a fully connected multi-layer NN in emulating the state-intensity dependency denoted by the non-linear function  $f$  in Equation (2). We randomly split  $\mathcal{D}$

into a training and a validation set (respectively, composed of the 90% and 10% of the  $\mathcal{D}$  data) and we train the NN for performing the regression task  $f : X \rightarrow I$ . Subsequently, 30 episodes of online iLQR are performed, in order to lead the FERMI intensity toward the 95% of the target  $I^*$  from each of the 30 previously selected initial states. At each  $k$ -th time instant of each episode, given the target augmented state  $\hat{x}^* = [x^*, I^*]^\top$ , the iLQR algorithm is fed by the current state  $x_{(k)}^R$  and intensity  $I_{(k)}^R$ , and outputs an optimal sequence  $\mathbf{u}^*$ , whose first element is applied to the system. Each episode ends either when  $I_{(k)} \geq 95\%I^*$ , or when the maximum number  $N_{\max}$  of control steps is exceeded. Figure 2 shows a schematic representation of the above described procedure.



**Figure 2.** Schematic representation of a single run of the online iLQR. The termination occurs when  $I_{(k)} \geq 95\%I^*$ , or when the control steps number exceeds  $N_{\max}$ .

In order to allow a comparison, each episode is also repeated (meaning that the same initial condition is used) performing a simple GA on the  $f$  function as approximated by the NN. Specifically, the control input at each  $k$ -th time instant is computed according to:

$$u_{(k)}^R = \zeta \cdot \frac{\nabla f(x_{(k)}^R)}{\|\nabla f(x_{(k)}^R)\|_2}, \tag{7}$$

where  $\zeta$  is the step size, and  $\nabla f$  is computed analytically from the network weights and structure, which are known. Each run ends once both the online iLQR and the GA have been performed for all the 30 initial states.

The experimental procedure is performed by allowing at most  $N_{\max} = 10$  steps for each episode, and by collecting  $N = 250$  additional samples at each run. Thus, the NN  $f$  used in  $\mathcal{S}$  is re-trained at each run with a data-set of increasing size ( $N_1 = 250, N_2 = 500, N_3 = 750, N_4 = 1000, N_5 = 1250, N_6 = 1500$ ), each time starting from null initial weights. The main idea is to find a minimum number of samples for the NN training that will bring the online iLQR approach to the 95% of the target intensity from each of the 30 starting conditions. We consider a fully connected multi-layer NN whose hyperparameters are shown in Table 1.

**Table 1.** Hyperparameters settings of the (fully connected) Multilayer Perceptron.

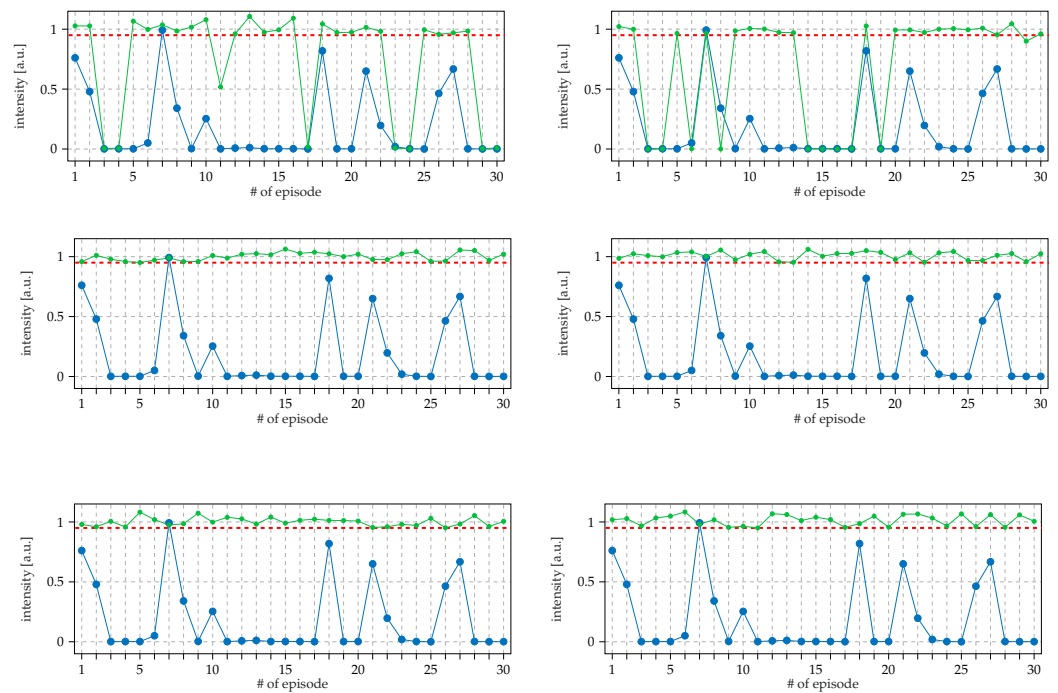
NN Hyperparameters			
# of hidden units (h.u.)	3		
# of neurons for h.u.	10	16	10
activation function for h.u.	tanh	sigmoid	sigmoid

The online iLQR procedure consists of LQR sub-problems of the form of Equation (4), with a horizon length  $H = 3$ , while  $Q = Q_H = \text{diag}\{1, 1, 1, 1, 1000\}$  and  $R = \text{diag}\{1, 1, 1, 1\}$  (where with  $\text{diag}\{a_1, a_2, \dots, a_n\}$  we denote a diagonal matrix whose diagonal entries are  $a_1, a_2, \dots, a_n$ ). The iLQR is performed imposing  $\text{tol} = 0.00001$  and a maximum number

of iterations equal to 25. The GA, conversely, allows fixed steps of size  $\zeta = 0.1$  along the maximization of the gradient.

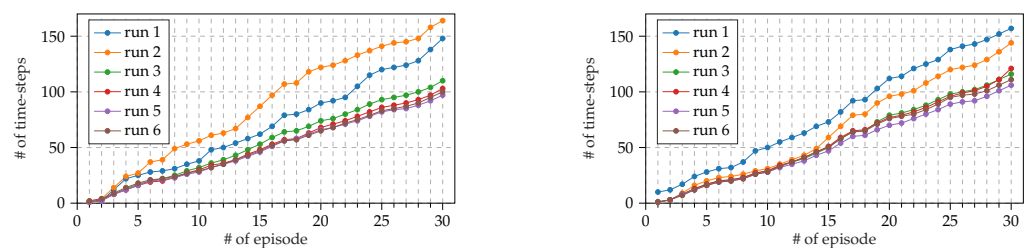
### 3.2. Results

Figure 3 shows the results obtained performing the online iLQR for each run and for each episode. In particular, each subplot reports the final intensity  $I_{(k)}^R$  (green) reached at each episode of a single run, starting from a particular initial intensity  $I_{(0)}^R$  (blue) and given the 95% of a predefined target intensity  $I^*$  (red). The results highlight the effectiveness of the proposed approach starting from the third run, where the NN is trained on a data-set composed of 750 samples, and the online iLQR reaches a final detected intensity greater than or equal to  $0.95I^*$  for all the episodes.

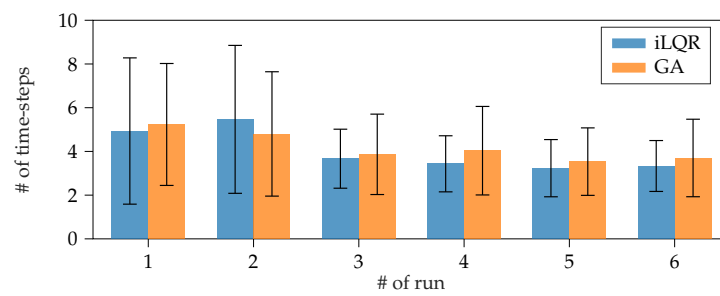


**Figure 3.** Normalized detected intensity at the end of each episode (green) for each online iLQR run, starting from an initial detected intensity (blue) and given a target intensity (red). The subfigures are run ordered, i.e., Section 3.2 reports the results of the first performed run, while Section 3.2 reports the results of the last performed run. Starting from the third run, the proposed approach results in achievement of the goal for each episode.

The same tests were conducted while also performing the control input resulting from a GA. Figures 4 and 5 compare the online iLQR and the GA in terms of the number of steps required to reach the target.



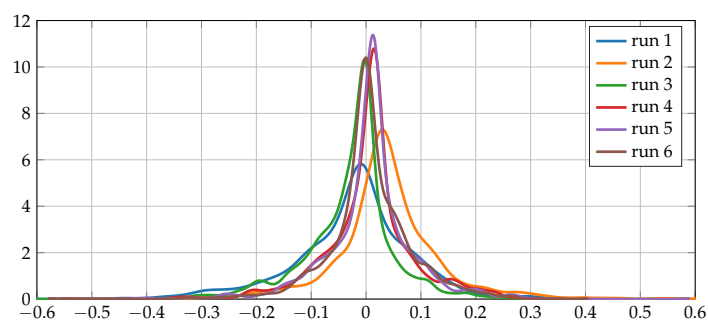
**Figure 4.** Cumulated time-steps during all the episodes for each run of the online iLQR (Section 3.2) and of GA (Section 3.2). Better performance is achieved by the online iLQR that requires a smaller total number of time-steps starting from the third test phase, the first completely successful one for both methods.



**Figure 5.** Statistical comparison of the number of steps required for the online iLQR and GA for the optimal working point attainment in FERMI. On average, the online iLQR performs the task in fewer steps.

The former shows the cumulative number of steps required for both methods in the achievement of the task for each episode and for each run. As evidenced by the last points of each subplot, the iLQR needs fewer steps than the GA, except in the first two runs, those in which the iLQR does not perform well (Figure 3). Figure 5, instead, reports the expected value and the variance of the number of steps required in each run by each of the two methods. Moreover, in this case it is apparent that the iLQR requires, on average, fewer steps (last two columns of Table 2).

A common feature of all the performed experiments is that, as the number of data provided for training the NN increases, the performance of both methods improves. This result is confirmed by Figure 6, showing the distribution of the error in the intensity estimation for each NN.



**Figure 6.** Distribution of the error in the intensity estimation ( $I_{(k)}^R - I_{(k)}$ ) of each neural network evaluated in a data-set composed of all the visited states during the performed episodes of each run.

For the purpose, we evaluated the NNs' capability of estimating the actual detected intensity in all the states visited during each episode of each run. The error distributions shrink their variance as the number of training data increases; moreover, they all present a mean value close to zero except for the first two, i.e., those trained with the two less numerous data-sets.

Table 2 summarizes the performance of both approaches in the different runs. It highlights that the iLQR reaches the 100% success rate starting from an NN with a 0.0072 of mean squared error in intensity estimation, while the GA needs a more accurate NN in order to produce a comparable performance.



**Table 2.** iLQR vs. GA. The best scores are highlighted in blue.

Run	NN MSE	Success Rate		Mean # of Steps	
		iLQR	GA	iLQR	GA
1	0.0133	73%	90%	4.93	5.2
2	0.0136	67%	87%	5.47	4.8
3	0.0076	100%	73%	3.67	3.87
4	0.0072	100%	97%	3.43	4.03
5	0.0069	100%	100%	3.23	3.53
6	0.0061	100%	100%	3.33	3.7

#### 4. Conclusions

We adapted the iLQR approach to the problem of beam alignment at FERMI. Specifically, we performed an identification step to obtain a computational model of the system (based on a NN), and we applied the iLQR in a closed-loop fashion. The proposed approach is generic, and thus it can potentially handle more complex dynamics and possible constraints on state and input (not considered here) [41]. We performed several experiments, with different amounts of data for training the NN. We have shown experimentally that the proposed approach is a viable one, and requires on average, fewer adjustment steps to reach an acceptable working point than a GA approach based on the same computational model. We neglected the time-varying nature of the system, and thus it has to be expected that, in the long run, the computational model needs to be updated. However, the employed NN has a simple structure and few weights, and thus a possible strategy could be retraining the network periodically by employing the data collected during the operation.

**Author Contributions:** Conceptualization, N.B., G.F., S.H., F.A.P. and E.S.; Validation, N.B., G.G. and S.H.; Formal analysis, N.B., G.F., F.A.P. and E.S.; Investigation, N.B., G.G. and S.H.; Writing—original draft preparation, N.B. and E.S.; Writing—review and editing, G.F., G.G., M.L. and F.A.P.; Supervision, M.L. and F.A.P. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** Not applicable.

**Conflicts of Interest:** The authors declare no conflict of interest.

#### Abbreviations

The following abbreviations are used in this manuscript:

FEL	Free-electron laser
SASE	Self-amplified spontaneous emission
SLAC	Stanford linear accelerator center
FLASH	Free electron laser in Hamburg
DESY	Deutsches elektronen-synchrotron
ML	Machine learning
RL	Reinforcement learning
FERMI	Free electron laser radiation for multidisciplinary investigation
iLQR	Iterative linear quadratic regulator
AWAKE	Advanced proton driven plasma wakefield acceleration experiment
CERN	Conseil européen pour la recherche nucléaire

NN	Neural network
MPC	Model-predictive control
GA	Gradient ascent
TT	Tip-tilt
YAG	Yttrium aluminum garnet
CCD	Charge-coupled device
LQR	Linear quadratic regulator

## References

- Colson, W. Theory of a free electron laser. *Phys. Lett. A* **1976**, *59*, 187–190. [[CrossRef](#)]
- Kim, K.J. An analysis of self-amplified spontaneous emission. *Nucl. Instrum. Methods Phys. Res. Sect. A Accel. Spectrom. Detect. Assoc. Equip.* **1986**, *250*, 396–403. [[CrossRef](#)]
- Yu, L.H. Generation of intense UV radiation by subharmonically seeded single-pass free-electron lasers. *Phys. Rev. A* **1991**, *44*, 5178. [[CrossRef](#)] [[PubMed](#)]
- Allaria, E.; Appio, R.; Badano, L.; Barletta, W.; Bassanese, S.; Biedron, S.; Borga, A.; Busetto, E.; Castronovo, D.; Cinquegrana, P.; et al. Highly coherent and stable pulses from the FERMI seeded free-electron laser in the extreme ultraviolet. *Nat. Photonics* **2012**, *6*, 699. [[CrossRef](#)]
- Allaria, E.; Castronovo, D.; Cinquegrana, P.; Craievich, P.; Dal Forno, M.; Danailov, M.; D’Auria, G.; Demidovich, A.; De Ninno, G.; Di Mitri, S.; et al. Two-stage seeded soft-X-ray free-electron laser. *Nat. Photonics* **2013**, *7*, 913. [[CrossRef](#)]
- Allaria, E.; Badano, L.; Bassanese, S.; Capotondi, F.; Castronovo, D.; Cinquegrana, P.; Danailov, M.; D’Auria, G.; Demidovich, A.; De Monte, R.; et al. The FERMI free-electron lasers. *J. Synchrotron Radiat.* **2015**, *22*, 485–491. [[CrossRef](#)] [[PubMed](#)]
- Tomin, S.; Geloni, G.; Zagorodnov, I.; Egger, A.; Colocho, W.; Valentinov, A.; Fomin, Y.; Agapov, I.; Cope, T.; Ratner, D.; et al. Progress in Automatic Software-based Optimization of Accelerator Performance. In Proceedings of the 7th International Particle Accelerator Conference (IPAC 2016), Busan, Korea, 8–13 May 2016.
- Bruchon, N.; Fenu, G.; Gaio, G.; Lonza, M.; Pellegrino, F.A.; Saule, L. Free-electron laser spectrum evaluation and automatic optimization. *Nucl. Instrum. Methods Phys. Res. Sect. A Accel. Spectrom. Detect. Assoc. Equip.* **2017**, *871*, 20–29. [[CrossRef](#)]
- Agapov, I.; Geloni, G.; Tomin, S.; Zagorodnov, I. OCELOT: A software framework for synchrotron light source and FEL studies. *Nucl. Instrum. Methods Phys. Res. Sect. A Accel. Spectrom. Detect. Assoc. Equip.* **2014**, *768*, 151–156. [[CrossRef](#)]
- McIntire, M.; Ratner, D.; Ermon, S. Sparse Gaussian Processes for Bayesian Optimization. In Proceedings of the Thirty-Second Conference on Uncertainty in Artificial Intelligence UAI, New York, NY, USA, 25–29 June 2016.
- McIntire, M.; Cope, T.; Ratner, D.; Ermon, S. Bayesian optimization of FEL performance at LCLS. In Proceedings of the 7th International Particle Accelerator Conference (IPAC 2016), Busan, Korea, 8–13 May 2016.
- Agapov, I.; Geloni, G.; Zagorodnov, I. Statistical optimization of FEL performance. In Proceedings of the 6th International Particle Accelerator Conference (IPAC 2015), Richmond, VA, USA, 3–8 May 2015.
- Radovic, A.; Williams, M.; Rousseau, D.; Kagan, M.; Bonacorsi, D.; Himmel, A.; Aurisano, A.; Terao, K.; Wongjirad, T. Machine learning at the energy and intensity frontiers of particle physics. *Nature* **2018**, *560*, 41–48. [[CrossRef](#)] [[PubMed](#)]
- Emma, C.; Edelen, A.; Hogan, M.; O’Shea, B.; White, G.; Yakimenko, V. Machine learning-based longitudinal phase space prediction of particle accelerators. *Phys. Rev. Accel. Beams* **2018**, *21*, 112802. [[CrossRef](#)]
- Edelen, A.; Neveu, N.; Frey, M.; Huber, Y.; Mayes, C.; Adelman, A. Machine learning for orders of magnitude speedup in multiobjective optimization of particle accelerator systems. *Phys. Rev. Accel. Beams* **2020**, *23*, 044601. [[CrossRef](#)]
- Fol, E.; de Portugal, J.C.; Franchetti, G.; Tomás, R. Optics corrections using Machine Learning in the LHC. In Proceedings of the 2019 International Particle Accelerator Conference, Melbourne, Australia, 19–24 May 2019.
- Azzopardi, G.; Salvachua, B.; Valentino, G.; Redaelli, S.; Muscat, A. Operational results on the fully automatic LHC collimator alignment. *Phys. Rev. Accel. Beams* **2019**, *22*, 093001. [[CrossRef](#)]
- Müller, R.; Balzer, A.; Baumgärtel, P.; Sauer, O.; Hartmann, G.; Viehhaus, J. Modernization of experimental data taking at BESSY II. In Proceedings of the 17th International Conference on Accelerator and Large Experimental Physics Control Systems, ICALEPCS2019, New York, NY, USA, 5–11 October 2019.
- Edelen, A.; Mayes, C.; Bowring, D.; Ratner, D.; Adelman, A.; Ischebeck, R.; Snuverink, J.; Agapov, I.; Kammering, R.; Edelen, J.; et al. Opportunities in machine learning for particle accelerators. *arXiv* **2018**, arXiv:1811.03172.
- Kain, V.; Hirlander, S.; Goddard, B.; Velotti, F.M.; Zevi Della Porta, G.; Bruchon, N.; Valentino, G. Sample-efficient reinforcement learning for CERN accelerator control. *Phys. Rev. Accel. Beams* **2020**, *23*, 124801. [[CrossRef](#)]
- Ramirez, L.V.; Mertens, T.; Mueller, R.; Viehhaus, J.; Hartmann, G. Adding Machine Learning to the Analysis and Optimization Toolsets at the Light Source BESSY II. In Proceedings of the 17th International Conference on Accelerator and Large Experimental Physics Control Systems, ICALEPCS2019, New York, NY, USA, 5–11 October 2019.
- Edelen, A.; Biedron, S.; Chase, B.; Edstrom, D.; Milton, S.; Stabile, P. Neural networks for modeling and control of particle accelerators. *IEEE Trans. Nucl. Sci.* **2016**, *63*, 878–897. [[CrossRef](#)]
- Edelen, A.L.; Edelen, J.P.; RadiaSoft, L.; Biedron, S.G.; Milton, S.V.; van der Slot, P.J. Using Neural Network Control Policies For Rapid Switching Between Beam Parameters in a Free-Electron Laser. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS), Long Beach, CA, USA, 4–9 December 2017.

24. Edelen, A.L.; Milton, S.V.; Biedron, S.G.; Edelen, J.P.; van der Slot, P.J.M. *Using A Neural Network Control Policy For Rapid Switching between Beam Parameters in an FEL*; Technical report; Los Alamos National Lab. (LANL): Los Alamos, NM, USA, 2017.
25. Hirlander, S.; Kain, V.; Schenk, M. New paradigms for tuning accelerators: Automatic performance optimization and first steps towards reinforcement learning at the CERN Low Energy Ion Ring. In *2nd ICFA Workshop on Machine Learning for Charged Particle*; 2019. Accelerators. Available online: [https://indico.cern.ch/event/784769/contributions/3265006/attachments/1807476/2950489/CO-technical-meeting-\\_Hirlander.pdf](https://indico.cern.ch/event/784769/contributions/3265006/attachments/1807476/2950489/CO-technical-meeting-_Hirlander.pdf) (accessed on 22 June 2021).
26. Bruchon, N.; Fenu, G.; Gaio, G.; Lonza, M.; Pellegrino, F.A.; Salvato, E. Toward the application of reinforcement learning to the intensity control of a seeded free-electron laser. In *Proceedings of the 2019 23rd International Conference on Mechatronics Technology (ICMT)*, Salerno, Italy, 23–26 October 2019; pp. 1–6.
27. Bruchon, N.; Fenu, G.; Gaio, G.; Lonza, M.; O’Shea, F.H.; Pellegrino, F.A.; Salvato, E. Basic reinforcement learning techniques to control the intensity of a seeded free-electron laser. *Electronics* **2020**, *9*, 781. [[CrossRef](#)]
28. O’Shea, F.; Bruchon, N.; Gaio, G. Policy gradient methods for free-electron laser and terahertz source optimization and stabilization at the FERMI free-electron laser at Elettra. *Phys. Rev. Accel. Beams* **2020**, *23*, 122802. [[CrossRef](#)]
29. Hirlander, S.; Bruchon, N. Model-free and Bayesian Ensembling Model-based Deep Reinforcement Learning for Particle Accelerator Control Demonstrated on the FERMI FEL. *arXiv* **2020**, arXiv:2012.09737.
30. Li, W.; Todorov, E. Iterative linear quadratic regulator design for nonlinear biological movement systems. In *Proceedings of the First International Conference on Informatics in Control, Automation and Robotics*, Setúbal, Portugal, 25–28 August 2004; pp. 222–229.
31. Kain, V.; Bruchon, N.; Goddard, B.; Hirlander, S.; Madysa, N.; Valentino, G.; Velotti, F. Sample-Efficient Reinforcement Learning for CERN Accelerator Control. The One World Charged Particle Accelerator (OWLE) Colloquium & Seminar Series. 2020. Available online: [https://drive.google.com/file/d/1-OcdIK57VDNznTOMkE\\_h28ZnTUqv7qza/view](https://drive.google.com/file/d/1-OcdIK57VDNznTOMkE_h28ZnTUqv7qza/view) (accessed on 22 June 2021).
32. Åström, K.J.; Wittenmark, B. *Computer-Controlled Systems: Theory and Design*; Courier Corporation: North Chelmsford, MA, USA, 2013.
33. Al-Duwaish, H.; Karim, M.N.; Chandrasekar, V. Use of multilayer feedforward neural networks in identification and control of Wiener model. *IEE Proc. Control Theory Appl.* **1996**, *143*, 255–258. [[CrossRef](#)]
34. Sjöberg, J.; Hjalmarsson, H.; Ljung, L. Neural networks in system identification. *IFAC Proc. Vol.* **1994**, *27*, 359–382. [[CrossRef](#)]
35. Nørgård, P.M.; Ravn, O.; Poulsen, N.K.; Hansen, L.K. *Neural Networks for Modelling and Control of Dynamic Systems—A Practitioner’s Handbook*. 2000. Available online: <https://orbit.dtu.dk/en/publications/neural-networks-for-modelling-and-control-of-dynamic-systems-a-pr> (accessed on 22 June 2021).
36. Jäntschi, L.; Bálint, D.; Bolboacă, S.D. Multiple linear regressions by maximizing the likelihood under assumption of generalized Gauss-Laplace distribution of the error. *Comput. Math. Methods Med.* **2016**, *2016*. [[CrossRef](#)] [[PubMed](#)]
37. Gómez, Y.M.; Gallardo, D.I.; Leão, J.; Gómez, H.W. Extended exponential regression model: Diagnostics and application to mineral data. *Symmetry* **2020**, *12*, 2042. [[CrossRef](#)]
38. Mzyk, G. Wiener System. In *Combined Parametric-Nonparametric Identification of Block-Oriented Systems*; Springer International Publishing: Cham, Switzerland, 2014; pp. 87–102. [[CrossRef](#)]
39. Lewis, F.L.; Vrabie, D.; Syrmos, V.L. *Optimal Control*; John Wiley & Sons: Hoboken, NJ, USA, 2012.
40. Borrelli, F.; Bemporad, A.; Morari, M. *Predictive Control for Linear and Hybrid Systems*; Cambridge University Press: Cambridge, UK, 2017.
41. Chen, J.; Zhan, W.; Tomizuka, M. Autonomous driving motion planning with constrained iterative LQR. *IEEE Trans. Intell. Veh.* **2019**, *4*, 244–254. [[CrossRef](#)]