On feedforward control using physics—guided neural networks: Training cost regularization and optimized initialization

Max Bolderman¹, Mircea Lazar¹, and Hans Butler^{1,2}

Abstract-Performance of model-based feedforward controllers is typically limited by the accuracy of the inverse system dynamics model. Physics-guided neural networks (PGNN), where a known physical model cooperates in parallel with a neural network, were recently proposed as a method to achieve high accuracy of the identified inverse dynamics. However, the flexible nature of neural networks can create overparameterization when employed in parallel with a physical model, which results in a parameter drift during training. This drift may result in parameters of the physical model not corresponding to their physical values, which increases vulnerability of the PGNN to operating conditions not present in the training data. To address this problem, this paper proposes a regularization method via identified physical parameters, in combination with an optimized training initialization that improves training convergence. The regularized PGNN framework is validated on a real-life industrial linear motor, where it delivers better tracking accuracy and extrapolation.

Index Terms—Neural networks, Feedforward control, Nonlinear system identification, Motion control, Linear motors.

I. Introduction

Inversion-based feedforward is a control method that improves reference tracking of dynamical systems by generating a control input based on a model of the inverse system dynamics, see, for example, [1]–[5]. Typically, linear or linear-in-the parameters physical models are used to parametrize the inverse system model [3], [6]. These types of model sets, however, are not able to approximate parasitic nonlinear forces that are always present in mechatronics, which significantly limits the achievable performance [7].

With the aim to approximate and compensate for nonlinear effects, neural networks (NN) have been originally proposed in [8] as inverse system parametrizations. NNs are known as universal function approximators [9], [10], and are therefore considered capable of identifying the complete inverse dynamics under suitable assumptions. Although, there exists some examples of successful implementation of NNs in feedforward control [11]–[13], the lack of physical interpretability and poor extrapolation capabilities [14] hinders adoption of NN–based controllers in industry.

To combine the benefits of physical models and neural networks, [15] introduced physics-guided neural networks

known, not necessarily linear, physics—guided part, with the output of an unknown, possibly nonlinear part that is parameterized by a black—box NN. Both the physics—guided and the NN part are embedded in a single model structure, which allows the PGNN to be identified as a single model. The considered model structure is in line with earlier parallel model structures utilized in nonlinear system identification, see, e.g., [16]–[18], which typically combine a nonlinear/NN model with a linear model. In [16], two standard approaches for identifying the parameters of such hybrid parallel models are distinguished:

1) Simultaneous identification of all parameters;

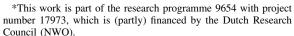
(PGNN) for feedforward control. The PGNN in [15] employs

a parallel model structure that combines the output of a

 Sequential (or consecutive) identification: first indentify the parameters of the physics-guided layer and then train the NN using the residuals of the physcsguided layer.

In principle, the first approach yields the best data fit and it was also used in [15] to train a feedforward controller. However, this approach can suffer from overaparameterization caused by the flexible nature of the NN that is capable of identifying also parts of the parallel physical model. This creates a parameter drift during training that may yield physically inconsistent parameter values for the physicsguided layer and poor extrapolation capabilities caused by the large NN outputs [14]. The second approach, used for example in [17], [18], preserves the physical model parameters in the physics-guided layer, which can be beneficial for extrapolation outside the training data set. However, approximation accuracy can be reduced because the training has less free parameters to optimize simultaneously. The sequential identification was shown in [16] to act as a regularization.

In this paper we propose the following contributions to the state of the art described above. First, we define a regularized cost function which penalizes, besides the standard data fit, the deviation of the parameters of the physics—guided layer with respect to some known/identified physical parameters. The known physical parameters can be obtained based on physical insights or via identification using a linear—in—the—parameters model (LIP). Second, we develop a novel initialization procedure for training of parallel PGNNs and we prove that it guarantees an improvement in accuracy of the data fitting with respect to the original LIP model. This is in line with recent initialization procedures for nonlinear system identification based on the best linear approximator,



¹Control Systems Group, department of Electrical Engineering, Eindhoven University of Technology, The Netherlands.

²ASML, Veldhoven, The Netherlands.

Email: m.bolderman@tue.nl, m.lazar@tue.nl, and h.butler@asml.nl



see [19] and the references therein. Experimental results on a coreless linear motor used in lithography industry confirm that the developed regularized training cost function and initialization procedure lead to a high performance and improved robustness to non-training data.

Remark 1.1: An alternative regularization method has been proposed in the physics-informed neural network (PINN) literature, see, e.g., [20]–[22]. These PINNs employ a black–box NN to identify an underlying function based on a cost function that penalizes, besides data fit, compliance of the NN output with the output of an available physical model. The penalty of deviating from the physical model requires the NN to optimize between either data fit or model fit, which results in a biased estimate when the underlying system dynamics deviates from the available physical model. Note that our PGNN design [15] differs by having the physical model as an intrinsic part of the PGNN, whereas the PINNs use the output of physical model only for regularization, i.e. the physical model itself is not embedded in the resulting PINN.

The remainder of this paper is organized as follows: Section II introduces the considered system dynamics and PGNN identification setup. The problem statement is formalized in Section III. Section IV describes the regularization—based PGNN cost function, and Section V explains the PGNN training initialization. Afterwards, the developed methodology is tested on an industrial coreless linear motor in Section VI, followed by conclusions in Section VII.

II. PRELIMINARIES

The symbols \mathbb{Z} and \mathbb{R} denote the set of integers and real numbers, respectively. The set $\mathbb{Z}_+ := \{i \in \mathbb{Z} \mid i > 0\}$ denotes the positive integers, and the same notation is used for positive real numbers, i.e., $\mathbb{R}_+ := \{i \in \mathbb{R} \mid i > 0\}$.

A. Inverse system dynamics and modelling

Consider the following discrete-time inverse dynamical system

$$u(t) = \theta_0^T T_{\text{phy}}(\phi(t)) + g(\phi(t)), \tag{1}$$

which describes the relation between the input u(t) and the output y(t), that corresponds to a nonlinear ARX model with a particular structure that is clarified as follows. In (1), $\phi(t) = [y(t+n_a), \ldots, y(t-n_b), u(t-1), \ldots, u(t-n_b)]^T$ is the regressor, with $n_a, n_b, n_c \in \mathbb{Z}_+$ the orders of the system, and $\theta_0 \in \mathbb{R}^{n_{\theta_0}}$ is the parameter vector corresponding to the known physical model. Furthermore, $T_{\text{phy}} : \mathbb{R}^{n_a+n_b+n_c+1} \to \mathbb{R}^{n_{\theta_0}}$ comprises of a set of functions that model the known part of the dynamics, e.g., based on physical insights, and $g : \mathbb{R}^{n_a+n_b+1} \to \mathbb{R}$ is an unknown nonlinear function that contains the *unknown dynamics*.

Within this paper, the following model structures are considered for identification of the inverse system dynamics (1).

Definition 1: A linear-in-the-parameters (LIP) model is defined as

$$\hat{u}(\theta_{\text{LIP}}, \phi(t)) = \theta_{\text{LIP}}^T T_{\text{phy}}(\phi(t)), \tag{2}$$

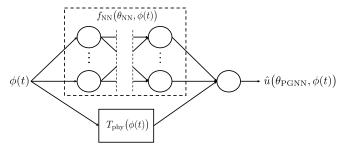


Fig. 1: Schematic overview of the considered PGNN structure (3), [15].

where $\hat{u}(\theta_{\text{LIP}}, \phi(t))$ is the predicted output of the model, and $\theta_{\text{LIP}} \in \mathbb{R}^{n\theta_0}$ are the parameters to be identified.

Definition 2: A physics-guided neural network (PGNN) is defined as

$$\hat{u}(\theta_{PGNN}, \phi(t)) = f_{NN}(\theta_{NN}, \phi(t)) + \theta_{phy}^{T} T_{phy}(\phi(t)), \quad (3)$$

where $\theta_{\text{PGNN}} = \{\theta_{\text{NN}}, \theta_{\text{phy}}\}$ are the PGNN parameters, with $\theta_{\text{NN}} = \{W_1, B_1, \dots, W_{l+1}, B_{l+1}\}$ the neural network weights and biases for each of the $l \in \mathbb{Z}_+$ hidden layers, and $\theta_{\text{phy}} \in \mathbb{R}^{n_{\theta_0}}$ are the parameters corresponding to the physics–guided part of the PGNN.

Remark 2.1: The NN output is recursively computed as

$$f_{\text{NN}}(\theta_{\text{NN}}, \phi(t)) = W_{l+1}\alpha_l(\dots \alpha_1(W_1\phi(t) + B_1)) + B_{l+1},$$
(4)

where α_i contains the activation functions that are applied element—wise.

The purpose of the PGNN model (3) is to exploit the NN part to identify the inverse dynamics (1) more accurately compared to the LIP model that is derived from physical knowledge (2). A visualization of the PGNN (3), see also [15], is shown in Figure 1.

B. Inverse system identification using the LIP model

Consider that an input-output data set

$$Z^{N} = \{u(0), y(0), \dots, u(N-1), y(N-1)\},$$
 (5)

is available that is generated by the actual system, i.e., Z^N satisfies the inverse system dynamics (1).

Definition 3: Two variables $x_a(t)$ and $x_b(t)$ are uncorrelated if and only if

$$\frac{1}{N}\sum_{t\in Z^N}x_a(t)x_b(t)=0. \tag{6}$$
 Let θ be a general vector of parameters, e.g., $\theta_{\rm LIP}$ or

Let θ be a general vector of parameters, e.g., θ_{LIP} or θ_{PGNN} . Then, identification is performed by choosing $\hat{\theta}$ as the minimizing argument of a cost function, i.e.,

$$\hat{\theta} = \arg\min_{\theta} V(\hat{u}(\theta, \phi(t)), Z^{N}). \tag{7}$$

Typically, the mean-squared error (MSE)

$$V(\hat{u}(\theta,\phi(t)),Z^N) = \frac{1}{N} \sum_{t \in Z^N} (u(t) - \hat{u}(\theta,\phi(t)))^2$$
 (8)

is chosen as a cost function. After identification of the parameters $\hat{\theta}$, the corresponding feedforward controller is obtained according to the following definition.

Definition 4: An inversion based feedforward controller is given as

$$u_{\rm ff}(t) = \hat{u}(\hat{\theta}, \phi_{\rm ff}(t)), \tag{9}$$

with $u_{\rm ff}(t)$ the feedforward signal, $\phi_{\rm ff} := [r(t+n_a), \dots, r(t-n_b), u_{\rm ff}(t-1), \dots, u_{\rm ff}(t-n_c)]^T$, and r(t) the reference.

Substitution of the LIP (2) in the identification criterion (7) with MSE cost function (8), gives the identified parameters

$$\hat{\theta}_{LIP} = M^{-1} \left[\frac{1}{N} \sum_{t \in Z^N} T_{phy} (\phi(t)) u(t) \right]. \tag{10}$$

where $M := \frac{1}{N} \sum_{t \in Z^N} T_{\text{phy}}(\phi(t)) T_{\text{phy}}(\phi(t))^T$. The solution (10) is unique if and only if M is non-singular.

Remark 2.2: In the situation that $T_{\text{phy}}(\phi(t)) = \phi(t)$, the identified parameters (10) becomes the best linear approximator (BLA) [19], [23].

Substitution of the LIP model (2) with identified parameters (10) into the inverse dynamics (1), gives

$$u(t) = \hat{\theta}_{LIP}^T T_{phy}(\phi(t)) + f(\phi(t)), \tag{11}$$

where $f(\phi(t)) := g(\phi(t)) + (\theta_0 - \hat{\theta}_{LIP})^T T_{phy}(\phi(t))$ denotes the *unmodelled dynamics*.

Lemma 2.1: The unknown dynamics $g\left(\phi(t)\right)$ are identical to the unmodelled dynamics $f\left(\phi(t)\right)$, if and only if $T_{\rm phv}\left(\phi(t)\right)$ and $g\left(\phi(t)\right)$ are uncorrelated.

Proof: The *unknown* and *unmodelled* dynamics are identical if and only if $\theta_0 - \hat{\theta}_{LIP} = 0$. Using the solution $\hat{\theta}_{LIP}$ (10) and the inverse dynamics (1) gives

$$\theta_{0} - \hat{\theta}_{LIP} = M^{-1}M\theta_{0} - M^{-1} \left[\frac{1}{N} \sum_{t \in Z^{N}} T_{phy}(\phi(t)) u(t) \right]$$
$$= M^{-1} \left[\frac{1}{N} \sum_{t \in Z^{N}} T_{phy}(\phi(t)) g(\phi(t)) \right], \tag{12}$$

which proves Lemma 2.1, since M is non-singular.

III. PROBLEM STATEMENT

A. Inverse system identification using the PGNN model

Identification of the inverse system dynamics (1) using the LIP model (2) leaves room for improvement, as there is still the unmodelled dynamics $f(\phi(t))$ to be identified. Therefore, it was proposed in [15] to use the PGNN model structure (3) for identification of the inverse system dynamics. Due to the parallel layer structure of the PGNN, if the cost function (8) is used for identification, there is no guarantee that the resulting parameters for the physical layer correspond to the parameters of the identified LIP model (2) of the original system (1). In fact, the NN part of the PGNN (3) can start to compete with the physical model.

Minimizing the cost (8) solely aims to optimize the data fit of the PGNN and does not attribute specific dynamics to the neural layer versus the physics—guided layer. It would be desirable to steer the training of the PGNN such that the physics–guided layer identifies the known physical dynamics and the NN layer identifies the remaining unmodelled dynamics. This is a well known problem in nonlinear system identification when models with a parallel structure are used, see [16], Chapter 21. Therein, a solution is proposed which is based on first identifying the known physical dynamics, and then training the NN using the residuals. This preserves the known dynamics in the physical part of the PGNN (3), but also limits achievable data fitting accuracy, as the whole set of parameters for the PGNN are not identified simultaneously.

B. Illustrative example of layer competition

In order to illustrate the competition between the NN and physics—guided layer in the PGNN (3), we consider the coreless linear motor (CLM) also used in [15]. A reasonably accurate LIP model can be obtained via first—principle modelling using Newton's second law, which gives

$$u(t) = \begin{bmatrix} m & f_v & f_c & f_k \end{bmatrix} \begin{bmatrix} \delta^2 y(t) \\ \delta y(t) \\ \operatorname{sign}(\delta y(t)) \\ y(t) \end{bmatrix} + g(\phi(t)), \quad (13)$$

where $\delta = \frac{1-q^{-1}}{T_s}$ is the backward Euler differentiation, with q^{-1} the backwards-shift operator, and T_s the sampling time. Also, $\phi(t) = [y(t), y(t-1), y(t-2)]^T$, and $m, f_v, f_c, f_k \in \mathbb{R}_+$ are the mass, viscous friction coefficient, Coulomb friction coefficient, and stiffness coefficient, respectively.

Identification using only the known, LIP model in (13) gives $\hat{\theta}_{\rm LIP} = [18.8, 172, 7.21, 1.36 \cdot 10^{-8}]^T$, whereas the corresponding parameters in the PGNN (13) converged to $\hat{\theta}_{\rm phy} = [18.2, 301, -4.99, 1.18 \cdot 10^3]^T$. The results are obtained using a single hidden layer NN with 16 neurons that have $\tanh(\cdot)$ activation functions. This parameter drift during training has two main disadvantages.

First of all, the parameter drift results in an inconsistent physics—guided layer in the PGNN (3). This complicates validation of the PGNN and thereby hinders applicability to safety critical systems. Secondly, extrapolation capabilities of the PGNN to operating conditions not present in the training data are severely limited. This is caused by the large output of the NN layer, and uninterpretable physics—guided layer, as is shown in Figure 2. Although the NN has universal approximation capabilities to identify the dynamics in the domain for which training data is generated, it is well known that its extrapolation capabilities are limited [14]. Therefore, it is desired to have the NN layer output small, while the parameters of the physics—guided layer of the PGNN remain close to the physical parameters.

C. Problem formulation

In order to solve the aforementioned problems, in this paper we propose a regularization–based approach for simultaneously training the PGNN (3). Under the condition that PGNN training converges to a global optimum, the regularized training allow for the following:

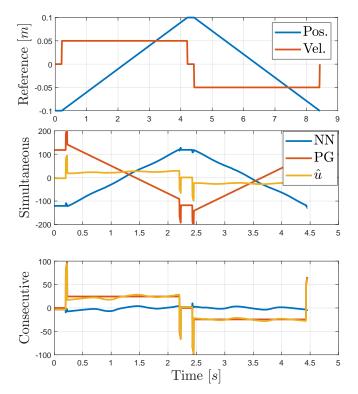


Fig. 2: Generated feedforward signals resulting from the reference (top), for a PGNN (3) where the NN and physics—guided (PG) layer are trained either simultaneously (middle), or sequentially (bottom).

- 1) When the NN layer can identify the *unmodelled dynamics*, the physics–guided layer must become the LIP, and the NN must identify the *unmodelled dynamics*.
- 2) When the NN cannot identify the *unmodelled dynamics*, the physics–guided layer parameters can still be partially used to improve data–fit.

Since training of the PGNN (3) is a non-convex optimization, it is possible that the optimization scheme ends up in a local minimum. Therefore, the second problem considered in this paper is defined as the design of a PGNN parameter initialization method that returns a smaller cost function compared to the LIP model. This property must also be satisfied for the regularized PGNN cost function.

IV. REGULARIZED PGNN TRAINING

In order to have the physical parameters θ_{phy} of the PGNN (3) remain close to the LIP parameters $\hat{\theta}_{\text{LIP}}$ (10), the regularized PGNN cost function is defined.

Definition 5: The regularized PGNN training cost function is defined as

$$V(\hat{u}(\theta_{PGNN}, \phi(t)), Z^{N}) = \frac{1}{N} \sum_{t \in Z^{N}} (u(t) - \hat{u}(\theta_{PGNN}, \phi(t)))^{2} + (\theta_{phy} - \hat{\theta}_{LIP})^{T} \Lambda(\theta_{phy} - \hat{\theta}_{LIP}),$$

$$(14)$$

where Λ is a symmetric positive definite matrix that defines the relative weights for penalizing deviation from the LIP.

Remark 4.1: The regularized PGNN training cost function (14) is fundamentally different from the PINNs regularization

$$V(\hat{u}(\theta_{NN}, \phi(t)), Z^{N}) = \frac{1}{N} \sum_{t \in Z^{N}} (u(t) - \hat{u}(\theta_{NN}, \phi(t)))^{2}$$

+ $\lambda \frac{1}{N} \sum_{t \in Z^{N}} (\hat{u}(\theta_{NN}, \phi(t)) - \theta_{phy}^{T} T_{phy}(\phi(t)))^{2},$ (15)

with $\lambda \in \mathbb{R}_+$ the relative weight of the regularization. Training of the PINNs can be done either by optimizing only $\theta_{\rm NN}$ and fixing $\theta_{\rm phy} = \hat{\theta}_{\rm LIP}$, or by simultaneously training $\{\theta_{\rm NN}, \theta_{\rm phy}\}$. Clearly, it is not possibly to have both contributions in (15) become zero when there is unknown dynamics, see $g(\phi(t))$ in (1). This implies that the PINNs approach [20]–[22] must optimize between either data fit or physical model compliance. Straightforwardly, this creates a bias in the estimation of $\hat{\theta}_{\rm NN}$.

It can be proven that the PGNN regularization (14) does not create a bias under the following assumptions.

Assumption 4.1: The NN in the PGNN (3) is chosen sufficiently rich to identify the unmodelled dynamics, i.e., there exists a θ_{NN}^* such that

$$f_{\text{NN}}(\theta_{\text{NN}}^*, \phi(t)) = f(\phi(t)). \tag{16}$$

This θ_{NN}^* is unique up until the level of interchanging neurons in the same layer.

Assumption 4.2: The data set \mathbb{Z}^N (5) is persistently exciting with respect to the NN part of the PGNN (3), such that

$$\frac{1}{N} \sum_{t \in Z^N} \left(f_{\text{NN}} \left(\theta_{\text{NN}}, \phi(t) \right) - f \left(\phi(t) \right) \right)^2 = 0 \implies \theta_{\text{NN}} = \theta_{\text{NN}}^*. \tag{17}$$

Proposition 4.1: Consider the data generating system (1), the PGNN (3), and the identification criterion (7) with regularized PGNN cost function (14). Suppose that Assumptions 4.1 and 4.2 hold, and let $\hat{\theta}_{\rm LIP}$ be calculated as in (10). Assume that the minimization of the cost function (14) converges to a global optimum. Then, the corresponding optimal parameters satisfy $\hat{\theta}_{\rm PGNN} = \{\hat{\theta}_{\rm LIP}, \theta_{\rm NN}^*\}$.

Proof: Writing out the regularized cost function (14) and substituting the PGNN (3) gives

$$V(\hat{u}(\theta_{PGNN}, \phi(t)), Z^{N}) = \frac{1}{N} \sum_{t \in Z^{N}} (u(t) - \theta_{phy}^{T} T_{phy}(\phi(t)) - f_{NN}(\theta_{NN}, \phi(t)))^{2} + (\theta_{phy}^{T} - \hat{\theta}_{LIP})^{T} \Lambda(\theta_{phy} - \hat{\theta}_{LIP}) \ge 0.$$

$$(18)$$

Both terms in (18) are non-negative. Correspondingly, the global optimum is attained if $\hat{\theta}_{phy} = \hat{\theta}_{LIP}$ (regularization term), and $\hat{\theta}_{NN} = \theta_{NN}^*$ (data-fit term, substitute u(t) from (1)).

Remark 4.2: It is in general difficult to know the required NN dimensions (number of hidden layers, number of neurons per hidden layer) for Assumption 4.1 to be satisfied. In the

situation that Assumption 4.1 is violated, the training of the PGNN (3) with regularized cost function (14) allows the physics—guided layer parameters to slightly deviate from the LIP in order to better fit the data.

V. PGNN TRAINING INITIALIZATION

Proposition 4.1 holds under the assumption that training of the PGNN (3) converges to the global minimum of the cost function (14). The non–convexity of the optimization however, requires the use of nonlinear optimization schemes that tend to get stuck in local minima, which strongly depends on the initialized parameters [24]. For this reason, there is a strong interest in finding initial parameter values that already perform well on the cost function (14).

In order to illustrate the proposed PGNN parameter initialization, we denote $\theta_{PGNN}^{(k)}$ as the value of θ_{PGNN} at the k'th iteration in training, and rewrite the PGNN output (3) as

$$\hat{u}(\theta_{PGNN}, \phi(t)) = W_{l+1}^T f_{HL}(\theta_{HL}, \phi(t)) + B_{l+1} + \theta_{phv}^T T_{phv}(\phi(t)),$$
(19)

where $\theta_{\rm HL}:=\{W_1,B_1,\ldots,W_l,B_l\}$, such that $\theta_{\rm NN}=\{\theta_{\rm HL},W_{l+1},B_{l+1}\}$. Also, $f_{\rm HL}:\mathbb{R}^{n_a+n_b+n_c+1}\to\mathbb{R}^{n_l}$ denotes the output of the last hidden layer $l\in\mathbb{Z}_+$ that contains $n_l\in\mathbb{Z}_+$ neurons. In contrary to the majority of NN literature that initializes all the weights and biases $\theta_{\rm NN}^{(0)}$ randomly, we only do so for the hidden layer weights and biases $\theta_{\rm HL}^{(0)}$, and define $\phi_{\rm OL}^{(0)}(t):=[f_{\rm HL}\big(\theta_{\rm HL}^{(0)},\phi(t)\big)^T,1,T_{\rm phy}\big(\phi(t)\big)^T]^T$. Then, under the following assumption, we prove that the PGNN can be initialized at a lower regularized cost function (14) compared to the LIP model (2).

Assumption 5.1: The data set \mathbb{Z}^N (5) is persistently exciting for the output weights and biases of the PGNN (3), i.e.,

$$M_{R} := \frac{1}{N} \sum_{t \in Z^{N}} \phi_{OL}^{(0)}(t) \phi_{OL}^{(0)}(t)^{T} + \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix} \Lambda \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix}^{T}$$
(20)

is non-singular.

Lemma 5.1: Consider the PGNN (3) with parameters $\theta_{\rm HL}^{(0)}$ initialized randomly, and the regularized cost function (14). Suppose that Assumption 5.1 holds. Define $\theta_{\rm OL}:=[W_{l+1}^T,B_{l+1},\theta_{\rm DN}^T]^T$ and initialize as

$$\theta_{\rm OL}^{(0)} = M_{\rm R}^{-1} \left[\frac{1}{N} \sum_{t \in Z^N} u(t) \phi_{\rm OL}^{(0)}(t) + \begin{bmatrix} 0 \\ 0 \\ I \end{bmatrix} \Lambda \hat{\theta}_{\rm LIP} \right]. \tag{21}$$

Then it holds that

$$V\left(\hat{u}\left(\hat{\theta}_{\text{PGNN}}^{(0)},\phi(t)\right),Z^{N}\right) \leq V\left(\hat{u}\left(\hat{\theta}_{\text{LIP}},\phi(t)\right),Z^{N}\right), \quad (22)$$

with strict inequality if and only if

$$M_{\mathbf{R}} \begin{bmatrix} 0\\0\\\hat{\theta}_{\mathrm{LIP}} \end{bmatrix} - \left(\frac{1}{N} \sum_{t \in Z^{N}} u(t) \phi_{\mathrm{OL}}^{(0)}(t) + \hat{\theta}_{\mathrm{LIP}}^{T} \Lambda \begin{bmatrix} 0\\0\\I \end{bmatrix}^{T} \right) \neq 0.$$
(23)

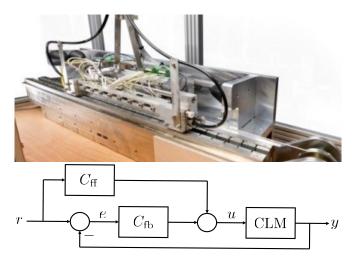


Fig. 3: Experimental CLM setup, and a schematic overview of the closed–loop control structure.

Proof: Rewriting the regularized cost function (14) gives

$$V(\theta_{\text{OL}}^{T}\phi_{\text{OL}}^{(0)}, Z^{N}) = \frac{1}{N} \sum_{t \in Z^{N}} u(t)^{2} + \hat{\theta}_{\text{LIP}}^{T}\Lambda\hat{\theta}_{\text{LIP}} + \theta_{\text{OL}}^{T}M_{\text{R}}\theta_{\text{OL}}$$
$$-2\left(\frac{1}{N} \sum_{t \in Z^{N}} u(t)\phi_{\text{OL}}^{(0)T} + \hat{\theta}_{\text{LIP}}^{T}\Lambda \begin{bmatrix} 0 & 0 & I \end{bmatrix} \right) \theta_{\text{OL}},$$

that has the global optimum $\theta_{OL}^{(0)}$ in (21), which is unique due to Assumption 5.1. Correspondingly, we can lowerbound the regularized cost function (14) according to

$$V\left(\hat{u}\left(\theta_{\text{PGNN}}^{(0)}, \phi(t)\right), Z^{N}\right) = V\left(\theta_{\text{OL}}^{(0)} \phi_{\text{OL}}^{(0)}(t), Z^{N}\right)$$

$$\leq V\left(\begin{bmatrix} 0 & 0 & \hat{\theta}_{\text{LIP}} \end{bmatrix}^{T} \phi_{\text{OL}}^{(0)}(t), Z^{N}\right)$$

$$= V\left(\hat{u}\left(\hat{\theta}_{\text{LIP}}, \phi(t)\right), Z^{N}\right).$$
(25)

In (25), the inequality holds with equality if and only if $\frac{\partial V\left(\theta_{\text{OL}}^{T}\phi_{\text{OL}}^{(0)}(t),Z^{N}\right)}{\partial\theta_{\text{OL}}}\Big|_{\theta_{\text{OL}}=[0,0,\hat{\theta}_{\text{LIP}}^{T}]^{T}}=0$. If (23) holds, this equality is not satisfied and we have (22) with strict inequality.

After initialization of the parameters $\theta_{PGNN}^{(0)} = \{\theta_{OL}^{(0)}, \theta_{HL}^{(0)}\}$ as in Lemma 5.1, training θ_{PGNN} is performed using a nonlinear optimization scheme that returns $\{\theta_{PGNN}^{(0)}, \ldots, \theta_{PGNN}^{(k)}\}$, with k the number of iterations for the solver to converge or stop. Then, choosing $\hat{\theta}_{PGNN}$ as the iteration with the smallest cost function, concludes that

$$V(\hat{u}(\hat{\theta}_{PGNN}, \phi(t)), Z^{N}) \leq V(\hat{u}(\theta_{PGNN}^{(0)}, \phi(t)), Z^{N})$$
$$< V(\hat{u}(\hat{\theta}_{LIP}, \phi(t)), Z^{N}), \tag{26}$$

if condition (23) holds.

Remark 5.1: Beneficial of the proposed PGNN (3) is that Lemma 5.1 does not depend on the dimensions of the NN part of the PGNN. This is in contrast to the PINNs approach, where there are no clear guidelines for the design of the NN.

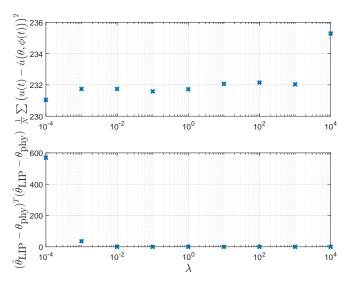


Fig. 4: Effect of the regularization $\Lambda = \lambda I$ (14) on the training convergence for the data fit (top) and LIP fit (bottom).

VI. EXPERIMENTAL VALIDATION

To illustrate the effectiveness of the regularized PGNN (3) for feedforward control as defined in (9), we consider the problem of position control for a real–life coreless linear motor (CLM) shown in Figure 3. The system is controlled in closed–loop by a feedback controller proposed in [15]. Training data is generated while operating the CLM in closed–loop at a frequency of 10~kHz, with a third order nominal reference $r_1(t) := \{r(0), \ldots, r(N_R-1)\}$ that moves back–and–forth in $r_1(t) \in \{-0.1, 0.1\}~m$, with maximum velocity $\max(|\dot{r}_1(t)|) = 0.05~\frac{m}{s}$, acceleration $\max(|\ddot{r}_1(t)|) = 4~\frac{m}{s^2}$, and jerk $\max(|\ddot{r}_1(t)|) = 1000~\frac{m}{s^3}$. Additionally, the input to the system u(t) is dithered with a normally distributed white noise $\Delta u(t) \sim \mathcal{N}(0, 50^2)$.

Following the previously introduced CLM dynamics (13), the PGNN is defined as

$$\hat{u}(\theta_{\text{PGNN}}, \phi_{\text{CLM}}(t)) = \theta_{\text{phy}}^T \phi_{\text{CLM}}(t) + f_{\text{NN}}(\theta_{\text{NN}}, \phi_{\text{CLM}}(t)),$$
(27)

with $\phi_{\text{CLM}} := [\delta^2 y(t), \delta y(t), \text{sign}(\delta y(t)), y(t)]^T$. The transformed inputs enter the NN, as this was shown to enhance convergence of the PGNN training for high sampling rates in [15]. Moreover, the NN part of the PGNNs (27) have a single hidden layer with $n_l = 16$ neurons that have $\tanh(\cdot)$ activation function.

Figure 4 shows the effect of the regularization using different values for $\Lambda = \lambda I$ on the training convergence for both the data fit and the LIP fit (14). There is a trend observable where a smaller λ allows the physics–guided parameters $\theta_{\rm phy}$ to deviate more from the $\hat{\theta}_{\rm LIP}$ in order to improve the data fit.

Figure 5 shows the tracking error e(t) := r(t) - y(t) resulting from the nominal reference $r_1(t)$ for the PGNNs (3) that are trained either sequentially, or with regularized cost function (14) using $\Lambda = \lambda I$ with $\lambda = 0$ or $\lambda = 0.01$. There is no major difference between the different PGNNs in terms

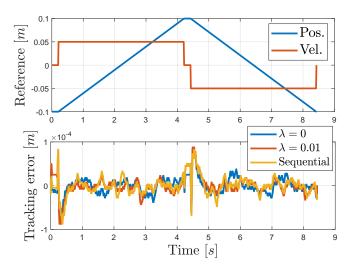


Fig. 5: Tracking error on the nominal reference $r_1(t)$ used for training data generation.

TABLE I: MAE of the tracking error e(t) for the nominal reference $r_1(t)$ and new reference $r_2(t)$.

		$r_1(t)$	$r_2(t)$
	No FF	$5.35 \cdot 10^{-5}$	$6.19 \cdot 10^{-5}$
	LIP	$2.53 \cdot 10^{-5}$	$2.64 \cdot 10^{-5}$
$n_l = 8$	$\lambda = 0$	$2.26 \cdot 10^{-5}$	$2.57 \cdot 10^{-5}$
	$\lambda = 0.01$	$1.84 \cdot 10^{-5}$	$2.47 \cdot 10^{-5}$
	Consecutive	$2.69 \cdot 10^{-5}$	$3.04 \cdot 10^{-5}$
$n_l = 16$	$\lambda = 0$	$1.39 \cdot 10^{-5}$	$3.36 \cdot 10^{-5}$
	$\lambda = 0.01$	$1.32 \cdot 10^{-5}$	$2.66 \cdot 10^{-5}$
	Consecutive	$1.38 \cdot 10^{-5}$	$2.87 \cdot 10^{-5}$

of tracking performance, as is also confirmed by the data in Table I that lists the mean-absolute error

$$MAE(e(t)) := \frac{1}{N_R} \sum_{t \in Z^R} |r(t) - y(t)|.$$
 (28)

Figure 6 shows the tracking error resulting from a reference $r_2(t)$ that oscillates between $r_2(t) \in \{0,0.17\}$ m, using the same bounds on velocity, acceleration, and jerk as $r_1(t)$. The regularization helps to keep performance loss limited when operating the PGNN feedforwards on conditions not present in the training data. This is also confirmed by the MAE shown in Table I. On the other hand, it is mainly due to the feedback controller that the MAE for $\lambda=0$ does not become larger on this experiment.

Additionally, Table I shows the results for $n_l=8$ hidden layer neurons. For this situation, in which the NN part of the PGNN has limited approximation capabilities, it is clear that the regularized PGNN significantly outperforms the sequentially (consecutively) trained PGNN. Currently, what limits the measurable tracking performance is the $0.5 \cdot 10^{-5}$ m resolution of the linear encoder with which position measurements y(t) are taken.

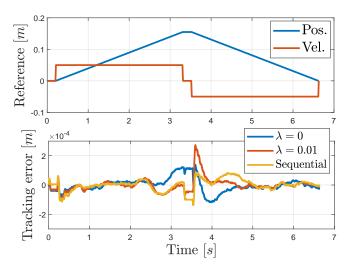


Fig. 6: Tracking error for the reference $r_2(t)$ that attains positions not present in the training data.

VII. CONCLUSION

In this paper, a regularization-based PGNN feedforward control framework was introduced for enabling a trade-off between high data fitting accuracy of NNs and the good extrapolation properties of physical models. An optimized initialization method was developed for the considered PGNN with a parellel structure, which has improved training convergence compared to a corresponding linear-in-the-parameter model. Experimental validation showed that developed regularization of the training cost is capable of optimizing the trade-off between data fit and physics-model based extrapolation.

REFERENCES

- M. L. G. Boerlage, M. Steinbuch, P. F. Lambrechts, and M. M. J. van de Wal, "Model-based feedforward for motion systems," *IEEE International Conference on Control Applications*, vol. 2, pp. 1158–1163, 2003.
- [2] J. A. Butterworth, L. Y. Pao, and D. Y. Abramovitch, "Analysis and comparison of three discrete-time feedforward model-inverse control techniques for nonminimum-phase systems," *Mechatronics*", vol. 22, no. 5, pp. 577–587, 2012.
- [3] Y. H. Yuen, M. Lazar, and H. Butler, "Data-driven neural feedforward controller design for industrial linear motors," 23rd International Conference on System Theory, Control and Computing, pp. 461–467, 2019.
- [4] J. van Zundert and T. Oomen, "On inversion-based approaches for feedforward and ILC," *Mechatronics*, vol. 50, pp. 282–291, 2018.
- [5] K. Igarashi, R. Igarashi, T. Atsumi, and S. Nakadai, "Feedforward control for track–seeking control in hard disk drive with sampled–data polynomial based on first–order hold," *IEEE International Conference* on Mechatronics, 2021.
- [6] M. Steinbuch and R. M. J. G. van de Molengraft, "Iterative learning control of industrial motion systems," *IFAC Proceedings Volumes*, vol. 33, no. 26, pp. 899–904, 2000.
- [7] S. Devasia, "Should model-based inverse inputs be used as feed-forward under plant uncertainty?" *IEEE Transactions on Automatic Control*, vol. 47, pp. 1865–1871, 2002.
- [8] O. Sørensen, "Additive feedforward control with neural networks," IFAC Proceedings Volumes, vol. 32, no. 2, pp. 1378–1383, 1999.
- [9] K. Hornik, M. B. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.

- [10] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Mathematics of Control, Signals and Systems*, vol. 2, pp. 303– 314, 1989.
- [11] G. Otten, T. J. A. de Vries, J. van Amerongen, A. M. Rankers, and E. W. Gaal, "Linear motor motion control using a learning feedforward controller," *IEEE Transactions on Mechatronics*, vol. 2, no. 3, pp. 179– 187, 1997.
- [12] X. Ren, Y. L. Chow, V. Venkataramanan, and F. L. Lewis, "Feedforward control based on neural networks for disturbance rejection in hard disk drives," *IET Control Theory & Applications*, vol. 3, no. 4, pp. 411–418, 2009.
- [13] G. Li, J. Na, D. P. Stoten, and X. Ren, "Adaptive neural network feedforward control for dynamically substructured systems," *IEEE Transactions on Control Systems Technology*, vol. 22, no. 3, pp. 944–954, 2014.
- [14] P. J. Haley and D. Soloway, "Extrapolation limitations of multilayer feedforward neural networks," in *Proceedings of Internation Joint* Conference on Neural Networks, vol. 4, pp. 25–30, 1992.
- [15] M. Bolderman, M. Lazar, and H. Butler, "Physics-guided neural networks for inversion-based feedforward control applied to linear motors," *IEEE Conference on Control Technology and Applications*, 2021
- [16] O. Nelles, Nonlinear System Identification. Springer, 2001.
- [17] T. Yasui, A. Moran, and M. Hayase, "Integration of linear systems and neural networks for identification and control of nonlinear systems," *IFAC Proceedings Volumes*, vol. 30, no. 11, pp. 783–788, 1997.
- [18] H. Zabiri, M. Ramasamy, T. D. Lemma, and A. Maulud, "Nonlinear system identification using integrated linear-nn models: serives vs. parallel structures," *International Conference on Modeling, Simulation, and Control*, vol. 10, pp. 33–37, 2011.
- [19] M. Schoukens and R. Tóth, "On the initialization of nonlinear LFR model identification with the best linear approximation," *IFAC Papers Online*, vol. 53, no. 2, pp. 310–315, 2020.
- [20] A. Karpatne, W. Watkins, J. Read, and V. Kumar, "Physics-guided neural networks (PGNN): An application in lake temperature modeling," arXiv preprint arXiv:1710.11431, 2017.
- [21] A. Daw, R. Q. Thomas, C. C. Carey, J. S. Read, A. P. Appling, and A. Karpatne, "Physics–guided architecture (PGA) of neural networks for quantifying uncertainty in lake temperature modeling." arXiv preprint arXiv:1911.02682, 2019.
- [22] M. Raissi, P. Perdikaris, and G. E. Karniadakis, "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations," *Journal of Computation Physics*, vol. 378, pp. 686–707, 2019.
- [23] R. Pintelon and J. Schoukens, System identification: A frequency domain approach. Wiley-IEEE Press, 2012.
- [24] J. Nocedal and S. Wright, *Numerical optimization*. Springer Science & Business Media, 2006.