Learning-based feedforward augmentation for steady state rejection of residual dynamics on a nanometer-accurate planar actuator system

Ioannis Proimadis[1]   JPROIMADIS@GMAIL.COM
Yorick Broens[2]   Y.L.C.BROENS@TUE.NL
Roland Tóth[2,3]   R.TOTH@TUE.NL
Hans Butler[2,4]   H.BUTLER@TUE.NL

Abstract

Growing demands in the semiconductor industry result in the need for enhanced performance of lithographic equipment. However, position tracking accuracy of high precision mechatronics is often limited by the presence of disturbance sources, which originate from unmodelled or unforeseen deterministic environmental effects. To negate the effects of these disturbances, a learning based feedforward controller is employed, where the underlying control policy is estimated from experimental data based on Gaussian Process regression. The proposed approach exploits the property of including prior knowledge on the expected steady state behaviour of residual dynamics in terms of kernel selection. Corresponding hyper-parameters are optimized using the maximization of the marginalized likelihood. Consequently, the learned function is employed as augmentation of the currently employed rigid body feedforward controller. The effectiveness of the augmentation is experimentally validated on a magnetically levitated planar motor stage. The results of this paper demonstrate the benefits and possibilities of machine-learning based approaches for compensation of static effects, which originate from residual dynamics, such that position tracking performance for moving-magnet planar motor actuators is improved.

Keywords: Gaussian Process, Motion control, Learning based feedforward

1. Introduction

In high-precision lithography, production of integrated circuits is realized by projecting extreme ultraviolet light on a silicon wafer using projection optics. In order to achieve high throughput and high reliability, the silicon wafer is positioned under the projection optics using the wafer stage module, which is a planar motor system that is capable of achieving nanometer accuracy in position tracking. The growing demands in the semiconductor industry result in the necessity to increase throughput, while still maintaining accurate positioning. In order to meet the throughput demands, highly aggressive acceleration profiles are required, which introduce high-frequent position tracking errors due to the limited stiffness of the mechanical structure. The currently implemented state-of-the-art planar stage configuration, which is based on the design proposed by Cho et al. (2001) and further discussed by Compter (2004), relies on a double stroke mechanism, where a magnetically levitated moving-coil motor is used for the coarse positioning of the mover. For fine positioning of the mover, the short stroke motor is used, which is actuated by voice coils.

As an alternative, planar motors based on a moving-magnet configuration have been investigated

---

[1] VDL ETG Technology & Development, Eindhoven 5651GH, The Netherlands
[2] Control Systems Group, Eindhoven University of Technology, 5600MB, The Netherlands
[3] Systems and Control Laboratory, Institute for Computer Science and Control, Kende u. 13-17, H-1111 Budapest, Hungary
[4] ASML, Veldhoven 5504DR, The Netherlands

in recent years, see de Boeij et al. (2009). In contrast with the moving-coil configuration, the moving-magnet configuration is comprised of a stator base and a freely floating magnet plate. The absence of physical connections between the mover and the environment results in a significant reduction of induced disturbances. Additionally, the moving-magnet configuration allows for a smaller and lighter moving body, see Proimadis (2020) and Rovers (2013), such that high accelerations can be achieved with relatively low power demands.

However, these advantages come at the cost of introducing additional complexity. Magnetically levitated planar motors exhibit complex non-linear multi-physical effects and are subject to various disturbances, which are machine specific in terms of the design of the magnet and coil arrays, see Rovers (2013). The highly complex dynamics can only be approximately modelled based on first principle knowledge and therefore the achievable position tracking performance is limited.

In order to achieve nanometer position tracking in magnetically levitated planar motor systems, feedforward control plays a crucial role, see Clayton et al. (2009). However, for accurate design of a feedforward controller, an accurate plant model is required, which, for moving-magnet planar actuator systems, is not trivial to obtain due to the complexity of the moving-magnet configuration. Therefore, standard feedforward control strategies do not provide the desired position tracking performance due to model mismatch between the first-principle based model and the real system.

In order to improve position tracking performance, learning based strategies can be employed to construct a feedforward policy, see Mooren et al. (2020), Goubej et al. (2019) and Proimadis (2020). By such a learning-based approach, the steady state behaviour of residual dynamics is captured by viewing it as a load disturbance. Then this load disturbance can be modelled using the Gaussian Process (GP) framework, which is advantageous since it ensures uncertainty bounds, such that reliability of the GP model is guaranteed.

Capturing the behaviour of unforeseen dynamics as a function of generalized coordinates allows for augmentation of the currently employed rigid body feedforward controller, such that static effects of residual dynamics of the magnetically levitated planar motor system are compensated for.

This paper is organized as follows. In Section 2, a brief description of the magnetically levitated planar motor system is presented. Section 3 presents the GP-based modelling of steady state behaviour of residual dynamics. Section 4 describes the experimental validation of the designed feedforward augmentation, where the experiments are performed on a magnetically levitated planar actuator. Lastly, in Section 5 the conclusions are drawn.

2. Magnetically levitated planar motor system

2.1. System overview

The Nanometer Accurate Planar Actuator System (NAPAS) prototype, which is based on a moving-magnet configuration, is illustrated in Figure 1. The NAPAS prototype consists of three separate coordinate frames: The stator base ($\mathcal{C}$), the translator ($\mathcal{T}$) and the metrology frame ($\mathcal{M}$). The stator base is a double layer coil array, consisting of 160 coils of which 40 coils are simultaneously activated at every time instant using 40 power converters, depending on the relative position of the translator. Proper actuation of the coils offers the means of both stabilization and propulsion of the magnet plate in 6 Degrees of Freedom (DOFs). The metrology frame, which rests on air mounts to suppress effects of any floor induced disturbances, is used as a global reference frame, such that position tracking accuracy can be evaluated. On the metrology frame, 9 laser interferometers (LIFMs) are mounted to measure the relative displacement of the translator with respect to the metrology frame. Additionally, two sets of eddy current sensors (ECS) are mounted as auxiliary

measurement systems, where one set of ECS is used for initialization of the NAPAS prototype and the second set of ECS is used to capture the displacements between the stator base and the metrology frame ($\mathcal{C}$-$\mathcal{M}$). In order to relate the coordinate frames to each other, rigid body coordinate frame transformations are applied, see Murray et al. (1994).
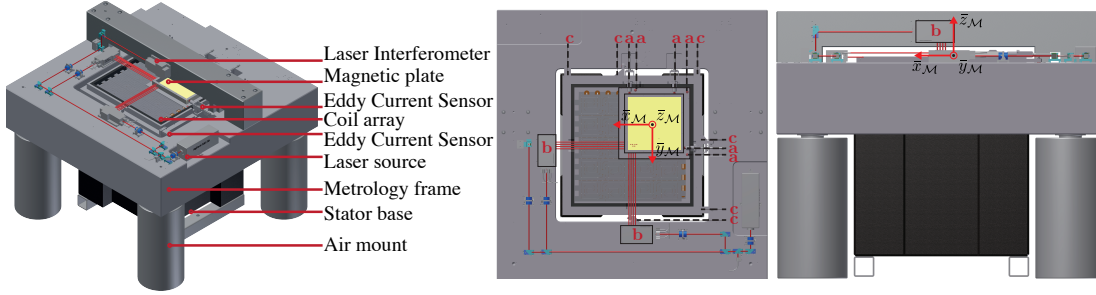

Figure 1: Schematic representation of the NAPAS setup, with a) ECSs used for initialization b) LIFM devices . c) ECS used to capture displacements between $\mathcal{C} - \mathcal{M}$. $\bar{x}_{\mathcal{M}}, \bar{y}_{\mathcal{M}}$ and $\bar{z}_{\mathcal{M}}$ denote the metrology coordinate frame.

The dynamic behaviour of the planar motor is governed by both electromagnetic and mechanical phenomena. The electromagnetic interaction describes the relation between the input currents in the coils and the resulting magnet force distribution on the magnet plate of the translator. Moreover, the mechanical model describes the relation between the aforementioned force distribution and the resulting motion of the translator.

2.2. Electromagnetic interaction

The actuation of the electromagnetically-levitated planar motor is realized by 40 power amplifiers, which independently control the current supplied to the corresponding 40 active coils. The relation between the supplied currents, $\boldsymbol{i} \in \mathbb{R}^{40}$, and the resulting force vector, is described by the Lorentz force principle (Rovers et al., 2012). Under the rigid body assumption, the force distribution, exerted on the magnet plate, is equivalently described as a force and torque vector around the center of mass (Murray et al., 1994). Since the position of the plate is evaluated with respect to the metrology frame, $\mathcal{M}$, it is convenient to define the force/torque vector with respect to the same frame. This force/torque vector is defined as

$$\boldsymbol{W}_{\mathcal{M}} = \begin{bmatrix} F_x & F_y & F_z & \tau_x & \tau_y & \tau_z \end{bmatrix}^{\top}, \tag{1}$$

where F_x and τ_x are the force and torque acting on the x axis (similarly for the other axes), expressed in the metrology frame. Then, the electromagnetic interaction is given by

$$\boldsymbol{W}_{\mathcal{M}} = R_{\mathcal{C}}^{\mathcal{M}} M^{\mathrm{r}} \left(\boldsymbol{q}_{\mathcal{T}}^{\mathcal{C}} \right) \boldsymbol{i}, \tag{2}$$

where $M^{\mathrm{r}} : \mathbb{R}^6 \mapsto \mathbb{R}^{6 \times 40}$ is a position dependent matrix, which depends on the relative position between the stator base and the mover, which is denoted as $\boldsymbol{q}_{\mathcal{T}}^{\mathcal{C}} = \begin{bmatrix} x_{\mathcal{M}}^{\mathcal{C}} & y_{\mathcal{M}}^{\mathcal{C}} & z_{\mathcal{M}}^{\mathcal{C}} & \chi_{\mathcal{M}}^{\mathcal{C}} & \psi_{\mathcal{M}}^{\mathcal{C}} & \zeta_{\mathcal{M}}^{\mathcal{C}} \end{bmatrix}^{\top}$. Moreover, since the force/torque vector is expressed with respect to the coil frame, the rotation matrix $R_{\mathcal{C}}^{\mathcal{M}}$ transforms the force and torque vector to the metrology coordinate frame.

2.3. Mechanical system

The kinematics of the translator are derived using Euler-Lagrange modelling strategies, see Jeltsema and Scherpen (2009). To linearise the dynamics of the translator, a small angle approximation is applied, since in lithographic equipment, rotational axes are actively steered to zero. Consequently, rotation angles are neglectable during operation, see Butler (2011). The resulting state space representation is expressed by (3), where $q_{\mathcal{T}}^{\mathcal{M}} = \begin{bmatrix} x_{\mathcal{T}}^{\mathcal{M}} & y_{\mathcal{T}}^{\mathcal{M}} & z_{\mathcal{T}}^{\mathcal{M}} & \chi_{\mathcal{T}}^{\mathcal{M}} & \psi_{\mathcal{T}}^{\mathcal{M}} & \zeta_{\mathcal{T}}^{\mathcal{M}} \end{bmatrix}^{\top}$ denotes the relative position of the mover with respect to the metrology frame and $G_{\mathcal{M}}$ denotes the gravity compensation vector.

$$\begin{bmatrix} \dot{q}_{\mathcal{T}}^{\mathcal{M}} \\ \ddot{q}_{\mathcal{T}}^{\mathcal{M}} \end{bmatrix} = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \begin{bmatrix} q_{\mathcal{T}}^{\mathcal{M}} \\ \dot{q}_{\mathcal{T}}^{\mathcal{M}} \end{bmatrix} + \begin{bmatrix} 0 \\ B \end{bmatrix} (W_{\mathcal{M}} - G_{\mathcal{M}}) \tag{3}$$

2.4. Control Principles

The control architecture of the NAPAS prototype, which is based on the standard motion control design for wafer scanners, is depicted in Figure 2. In order to apply SISO control strategies for first-principles feedback control design, the system is rigid body decoupled (Steinbuch et al. (2010)), where the commutation $\hat{\Gamma}(q_{\mathcal{T}}^{\mathcal{C}}) = \begin{bmatrix} R_{\mathcal{C}}^{\mathcal{M}} M^{\mathrm{r}}(q_{\mathcal{T}}^{\mathcal{C}}) \end{bmatrix}^{\dagger}$ denotes the actuator decoupling (mapping from control forces to required currents, $\hat{W}_{\mathcal{M}} \in \mathbb{R}^6 \mapsto i \in \mathbb{R}^{40}$) and the state reconstruction $\Psi(q_{\mathcal{T}}^{\mathcal{M}})$ denotes the sensor decoupling (mapping from independent LIFM measurements to the estimated physical axes, $\mathbf{y}_{\mathrm{L}} \in \mathbb{R}^9 \mapsto \hat{q}_{\mathcal{T}}^{\mathcal{M}} \in \mathbb{R}^6$). By applying rigid body decoupling strategies, the equivalent plant, as seen by the feedback controller, is expected to be diagonal up until the desired bandwidth (van de Wal et al., 2002).

For construction of a first-principle model-based feedback controller, a combination of a PID controller and a low-pass filter is considered for all six physical axes. Additionally, a standard rigid body feedforward is employed to the control system for further enhancement of position tracking performance.

Since the feedback controller, including the rigid body decoupling matrices, is based on the first-principles model, position tracking performance is limited by the neglected complex electromechanic phenomena which are present in the real system, but are very difficult to model from a first-principle view point. In order to improve position tracking performance, the steady state solution of the residual dynamics is viewed as load disturbance, which is modelled using the GP framework. Consequently, the control policy to compensate for the neglected dynamics is estimated from experimental data using Gaussian regression, which allows for augmentation of the identified GP model in the currently employed rigid body feedforward controller as illustrated in Figure 2.
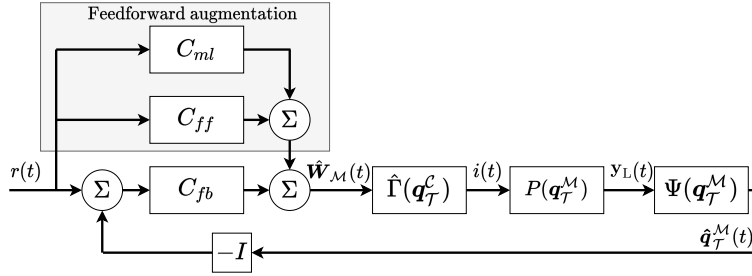

Figure 2: Schematic representation of the control architecture of the NAPAS prototype.

3. Gaussian-Process based learning of disturbances

3.1. Disturbance characteristics and experiment design

This section describes the modelling of static behaviour of residual dynamics using the GP framework. The main focus of this work is directed towards the suppression of disturbances acting on the z axis, which exhibits the largest deviations from the desired position in the investigated planar motor due to effects of residual dynamics, see (Proimadis, 2020). Apart from cancelling the known non-linear effects associated with the rigid-body dynamics, the commutation algorithm, presented in Section 2, further facilitated the control design by decoupling the system and reducing the number of inputs from 40, i.e. the number of active coils, to 6, i.e. the rigid-body of force and torque components. For similar reasons, it is desirable to view the static effect of residual dynamics acting on the z axis as load disturbances acting on the output $\hat{F}_z$ of the RB decoupled feedback controller. Using the Gaussian Process modelling framework, the disturbance force model is described by

$$y(i) = F_z^{\text{dist}}\left(\boldsymbol{w}\left(i\right)\right) + e(i), \tag{4}$$

where i is the time index and e is the noise source, modelled as i.i.d. white Gaussian noise, $e(i) \sim \mathcal{N}\left(0, \sigma_e^2\right)$ (Dekking et al., 2005), while $y \in \mathbb{R}$ is the measured output. The disturbance force F_z^{dist} is assumed to be dependent on the input vector $\boldsymbol{w} \in \mathbb{R}^{n_u}$ and will be modelled by a Gaussian Process (Rasmussen and Williams, 2005),

$$F_z^{\text{dist}}(\boldsymbol{w}) \sim \mathcal{GP}\big(0, k\left(\boldsymbol{w}(i), \boldsymbol{w}(j)\right)\big), \tag{5}$$

i.e. its covariance is fully characterized by the kernel function k. The Gaussian Process framework exploits the assumption of Gaussian distribution of both the function to be estimated and the involved noise process, in order to arrive at an analytically computed predictor, which is based on the posterior predictive distribution of F_z^{dist} conditioned on a training set of input-output observations. The mean of the predictive distribution is also the maximum a posteriori estimate

$$\bar{F}_z^{\text{dist}}\left(\boldsymbol{w}^*\right) = \boldsymbol{k}\left(\boldsymbol{w}^*, W_N\right)\left(\mathcal{K}\left(W_N, W_N\right) + \sigma_e^2 I_N\right)^{-1}\boldsymbol{y}, \tag{6}$$

where $\boldsymbol{w}^* \in \mathbb{R}^{n_u}$ is a test input, and W_N with $\boldsymbol{y}$ contain the training input and output data, respectively, at the N training points. For the vector $\boldsymbol{k}\left(\boldsymbol{w}^*, W_N\right) \in \mathbb{R}^{1 \times N}$, its ith element is equal to $k\left(\boldsymbol{w}^*, \boldsymbol{w}\left(i\right)\right)$, and in a similar fashion, for the matrix $\mathcal{K}\left(W_N, W_N\right) \in \mathbb{R}^{N \times N}$ the element on the ith row and j^{th} column is equal to $k\left(\boldsymbol{w}\left(i\right), \boldsymbol{w}\left(j\right)\right)$.

Based on the discussion so far, three main questions have to be answered towards the modelling of the disturbance via the Gaussian Process framework. First, in spite of the model postulated in (4), the static effects of the residual dynamics cannot be directly measured. Consequently, we need a methodology to infer the non-measurable values of these static effects from measurable data. Secondly, the relevant inputs have to be defined. Thirdly, the kernel function has to be chosen.

In order to answer the aforementioned questions, in Figure 3, the total force command on z axis on a selected grid of positions in the $x_{\mathcal{T}}^{\mathcal{M}} - y_{\mathcal{T}}^{\mathcal{M}}$ plane is shown, after the dynamic behaviour has settled. The two plots correspond to two independent measurements. At steady state, due to the integral action in the feedback controller, the resulting total control effort in each DOF is equal to the feedforward effort, minus the constant disturbances that act on the same DOF, i.e.

$$\hat{F}_z = \hat{F}_z^{\text{FF}} - F_z^{\text{dist}}, \tag{7}$$

with $\hat{F}_z^{\text{FF}}$ denoting the feedforward effort. By making use of (7), it becomes apparent that the measurements in Figure 3 can be directly used to estimate the magnitude of the disturbance.
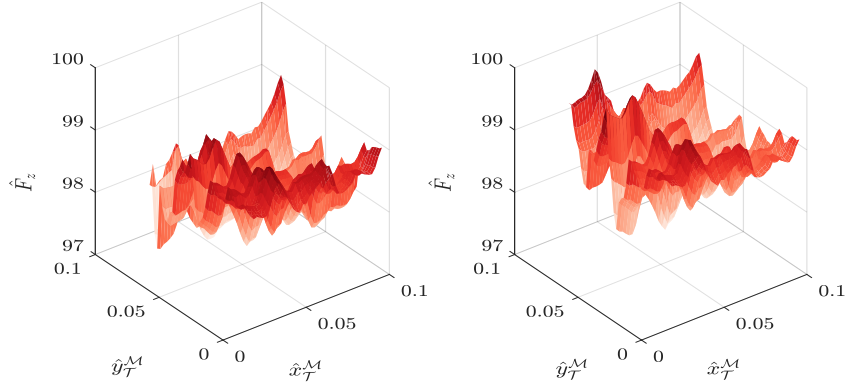
Figure 3: Two different measurements of $\hat{F}_z$ on the $\hat{x}_{\mathcal{T}}^{\mathcal{M}} - \hat{y}_{\mathcal{T}}^{\mathcal{M}}$ plane.

Figure 3 can be used further to draw conclusions, which are important for inferring which signals are affecting the static behaviour of residual dynamics, which are viewed as load disturbances. As a starting point, due to the position-dependent characteristics of the planar motor behaviour, it is natural to assume that the disturbances exhibit position-dependent characteristics as well. Secondly, it is asserted that the load disturbance exhibits a sinusoidal behaviour over the $x_{\mathcal{T}}^{\mathcal{M}} - y_{\mathcal{T}}^{\mathcal{M}}$ plane. This is verified by computing the 2D Fourier transform of the measured $\hat{F}_z$ with respect to the latter coordinates, shown in Figure 4, where it becomes evident that a sinusoidal behaviour with a spatial frequency of approximately $2.5 - 3$ cm is dominant. This value is related to the expected magnet pitch, see Custers (2019). Thirdly, the differences in the two measurements, presented in Figure 3, reveal that the residual dynamics cannot be solely described as a function of $\hat{x}_{\mathcal{T}}^{\mathcal{M}}, \hat{y}_{\mathcal{T}}^{\mathcal{M}}$. The deviation per grid point is on average equal to 0.46N, which is approximately equal to 0.5% of the total mass feedforward control force $\hat{F}_z^{\text{FF}}$. Equation (2) reveals a plausible source of residual dynamics; the vector $\hat{W}_{\mathcal{M}}$ depends on the relative position between the coil, metrology and translator frame. Consequently, any unforeseen, neglected or inaccurately modelled dynamics will depend, to some extent, on the relative distance between the aforementioned frames. Consequently, at least six input variables are required in order to relate the coordinate frames to each other.
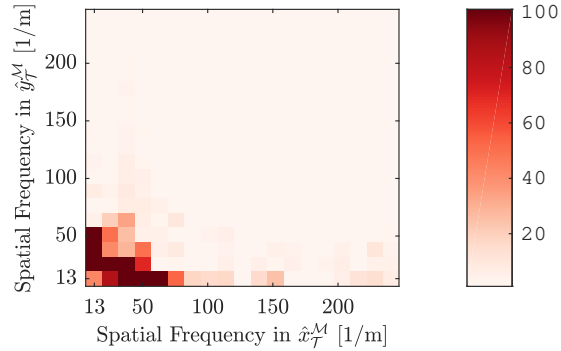

Figure 4: Spatial Fourier transform with respect to the $\hat{x}_{\mathcal{T}}^{\mathcal{M}}, \hat{y}_{\mathcal{T}}^{\mathcal{M}}$ coordinates.

The attenuation capabilities of the resulting feedforward control depend on how well the input space has been explored. The variables $\hat{q}_{\mathcal{T}}^{\mathcal{M}}$ are actively controlled, therefore it is straightforward

6

to define the excitation region with respect to them. The experiments have been performed on a $\hat{x}_{\mathcal{T}}^{\mathcal{M}} - \hat{y}_{\mathcal{T}}^{\mathcal{M}}$ grid, with $\hat{x}_{\mathcal{T}}^{\mathcal{M}}, \hat{y}_{\mathcal{T}}^{\mathcal{M}} \in [0.01, 0.1]$ m. The position difference between two consecutive measurement points is approximately equal to 2 mm. On the other hand, the $\hat{z}_{\mathcal{T}}^{\mathcal{M}}$ as well as the rotation of the translator with respect to the metrology frame are kept constant during operation. Regarding the $\boldsymbol{q}_{\mathcal{M}}^{\mathcal{C}}$ coordinates, the mounting of the metrology frame on passively controlled air mounts, as explained in Section 2, means that the relative displacement between the coil and the metrology frame cannot be actively controlled. Hence, the input space exploration is limited to the variability of the experiments, which is affected by environmental or other, unforeseen physical phenomena. For this reason, the experiment on the aforementioned $\hat{x}_{\mathcal{T}}^{\mathcal{M}}, \hat{y}_{\mathcal{T}}^{\mathcal{M}}$ grid has been repeated six times and the corresponding $\hat{\boldsymbol{q}}_{\mathcal{M}}^{\mathcal{C}}$ have been estimated. In total, the input vector is defined as

$$\boldsymbol{w}(i) = \begin{bmatrix} \hat{x}_{\mathcal{T}}^{\mathcal{M}}(i) & \hat{y}_{\mathcal{T}}^{\mathcal{M}}(i) & \hat{x}_{\mathcal{M}}^{\mathcal{C}}(i) & \hat{y}_{\mathcal{M}}^{\mathcal{C}}(i) & \hat{z}_{\mathcal{M}}^{\mathcal{C}}(i) & \hat{\chi}_{\mathcal{M}}^{\mathcal{C}}(i) & \hat{\psi}_{\mathcal{M}}^{\mathcal{C}}(i) & \hat{\zeta}_{\mathcal{M}}^{\mathcal{C}}(i) \end{bmatrix}^{\top}, \qquad (8)$$

with $\boldsymbol{w}(i) \in \mathbb{R}^8$ and $i = \{1, \ldots, N_p\}$, where N_p denotes the number of data points. Moreover, $w_\nu(i)$ denotes the νth element of the vector. Based on these inputs, an appropriate kernel selection is sought next.

3.2. Kernel design and validation

In the Gaussian Process framework, the kernel structure constitutes an important aspect, since it specifies the hypothesis space for the resulting estimator. As such, it is beneficial to incorporate any observations to specify the structure of the expected function class. To this end, the aforementioned load disturbance characteristics can be inscribed in the kernel structure, thus facilitating the fitting of an appropriate GP model to the experimental data. Based on the experimental observations in Section 3.1, the kernel is defined as the product of a periodic kernel (MacKay, 1998), which captures the coil array characteristics, a Radial Basis Function (RBF) kernel (Neal, 2012), which captures inconsistencies in the magnetic force distribution and a linear kernel, which explores linear trends with respect to the $\hat{x}_{\mathcal{T}}^{\mathcal{M}}$-$\hat{y}_{\mathcal{T}}^{\mathcal{M}}$ frame, which would cause $\hat{z}_{\mathcal{T}}^{\mathcal{M}}$ to vary as a function of $\hat{x}_{\mathcal{T}}^{\mathcal{M}}, \hat{y}_{\mathcal{T}}^{\mathcal{M}}$. The resulting kernel is mathematically described by

$$k(\boldsymbol{w}(i), \boldsymbol{w}(j)) = \sigma_1^2 \exp\Big(- \overbrace{\sum_{\nu=1}^{8} \frac{(w_\nu(i) - w_\nu(j))^2}{2\lambda_{\nu,\text{rbf}}^2}}^{\text{RBF kernel}} - \overbrace{\sum_{\nu=1}^{8} \frac{\sin^2\left(\frac{\pi(w_\nu(i)-w_\nu(j))}{p_{\sin}}\right)}{2\lambda_{\sin}^2}}^{\text{Periodic kernel}} \Big) +$$
$$\underbrace{\sum_{\nu=1}^{2} \sigma_{\nu,2}^2 (w_\nu(i) - c_{\nu,\text{lin}})(w_\nu(j) - c_{\nu,\text{lin}})}_{\text{Linear kernel}}, \qquad (9)$$

where $\sigma_{\nu,2}, c_{\nu,\text{lin}}$ is the νth element of the vector $\boldsymbol{\sigma}_2, \boldsymbol{c}_{\text{lin}}$ respectively. In total, there are 16 hyperparameters to be optimized, including the noise variance σ_e. Additionally, variations of the kernel structure, given by (9) have been investigated. More specifically, the kernel structure in (9) without the sinusoidal term, i.e. linear plus RBF, has been investigated, as well as a standalone RBF kernel. These two kernel structures are comprised of 14 and 10 hyperparameters, respectively. For all the aforementioned kernels, the hyperparameters have been computed by maximizing the marginal likelihood with respect to the hyperparameters (Rasmussen and Williams, 2005), using the collected training set.

For the validation of the results, 5 data sets with sufficient variations in $\boldsymbol{q}_{\mathcal{M}}^{\mathcal{C}}$ are used for training

purposes, where in total 3600 training points have been used. The predicted output $\hat{\boldsymbol{y}}$ is compared to the measured output $\boldsymbol{y}$ on a fresh data set using the Best Fit Ratio (BFR)

$$\text{BFR} = 100\% \cdot \max\left(1 - \frac{||\boldsymbol{y} - \hat{\boldsymbol{y}}||_2}{||\boldsymbol{y} - \bar{y}||_2}, 0\right),\tag{10}$$

where $\bar{y}$ denotes the sample mean of $\boldsymbol{y}$. The BFR criterion delivers an estimate between 0% for no match and 100% for the perfect match between the estimated and the measured output. The corresponding BFR results are presented in Table 1.

Table 1: Validation of the estimated GP compensator on the validation data set

Kernel	Linear + RBF $\times$ Sinusoidal	Linear + RBF	RBF
BFR %	85.77	77.60	77.62

A few conclusions can be drawn from Table 1. First of all, it is evident that the employed kernels manage to efficiently capture the major characteristics of the disturbances that act on the motor prototype. Among the three kernels, the kernel containing the sinusoidal term leads to the highest accuracy, thus verifying the physics-based intuition for selecting such a kernel. However, the remaining two kernels also manage to predict the steady state behaviour of residual dynamics in a satisfactory manner, which highlights the flexibility of the RBF kernel in capturing non-linear trends. Finally, it is observed that any linear trend can be captured by the RBF kernel alone, thus explaining the similarity in results for the Linear+RBF and RBF kernels. In total, the results show that the GP framework is a powerful tool for modelling highly complex, position-dependent static effects of residual dynamics, which can severely limit the positioning accuracy of planar motors.

4. Experimental results

In this section, the GP-based predictor is used for compensation of static disturbances, which are induced by residual dynamics that the feedback controller must compensate. Real-time implementation of the feedforward augmentation on the experimental prototype requires additional consideration. First of all, for $\hat{x}_{\mathcal{T}}^{\mathcal{M}}$, $\hat{y}_{\mathcal{T}}^{\mathcal{M}}$, the corresponding reference signals are used in the predictor instead of their measurement-based estimates. This selection is justified by the fact that for these two controlled variables, the error in this high-precision motor is relatively small compared to the sensitivity of the GP with respect to these variables. Moreover, introduction of additional feedback control loops is avoided, which could otherwise endanger stability. Additionally, the total control effort, shown in Figure 3, contains the feedforward control effort, i.e. the gravity compensation term. Therefore, in order to only compensate for the disturbance, the gravity compensation term is subtracted from the predicted disturbance force.

In order to evaluate the attenuation capabilities of the proposed feedforward controller in the NA-PAS prototype, its performance is tested under both diagonal as well as straight motion profiles with respect to the $\hat{x}_{\mathcal{T}}^{\mathcal{M}}$-$\hat{y}_{\mathcal{T}}^{\mathcal{M}}$ plane. For experimental validation of the designed feedforward augmentation, a low bandwidth feedback controller (≈ 9.5 Hz) is applied to obtain a better visualization of the effects of the designed feedforward augmentation on the position tracking performance. The selected plane is defined as $\hat{x}_{\mathcal{T}}^{\mathcal{M}}, \hat{y}_{\mathcal{T}}^{\mathcal{M}} \in [0.015, 0.055]$m. The employed 4th order reference trajectory is generated based on the method by Lambrechts et al. (2005). The designed reference trajectory is

illustrated in Figure 5.

For the real-time implementation of the GP-based feedforward augmentation, the computational complexity of the predictor has to be taken into consideration. For demonstration purposes, the linear plus RBF kernel is implemented in the experimental prototype, which, in practice, offers a good balance between prediction accuracy and required real-time calculation time. To further reduce the computation time, the initial predictor, which is formulated using the 3600 training points, is further approximated using a subset of 200 points. To accomplish this approximation, the Subset of Regressors (SR) (Wahba, 1990) method is applied, for which is empirically shown (Rasmussen and Williams, 2005, Chapter 8) that it offers a good balance between approximation error and computation time. For the selection of the subset of training points, a heuristic approach is considered (Proimadis, 2020). First, 200 samples are randomly selected for which the corresponding kernel is constructed. Secondly, the prediction capability of the resulting predictor is validated on a fresh data set by computing the BFR value. The heuristic approach is repeated 1000 times, after which the predictor with the highest BFR is kept. Using this approach, the approximated predictor achieved a BFR value of 65.37%, compared to the original BFR value of 77.60%.

The experimental results are shown in Figure 5. The trajectory consists of a motion on the $+y_{\mathcal{T}}^{\mathcal{M}}$ direction, followed by a diagonal motion on the $+x_{\mathcal{T}}^{\mathcal{M}}, -y_{\mathcal{T}}^{\mathcal{M}}$ direction, another $+y_{\mathcal{T}}^{\mathcal{M}}$ motion and finally the magnet plate returns to the starting point. Moreover, the error on the $\hat{z}_{\mathcal{T}}^{\mathcal{M}}$ axis is plotted for two cases, namely when the GP-based feedforward augmentation is active and inactive. Due to the coupling between the various DOFs, it is observed that the motion in the $x_{\mathcal{T}}^{\mathcal{M}}$ and/or $y_{\mathcal{T}}^{\mathcal{M}}$ axes leads to higher error in $\hat{z}_{\mathcal{T}}^{\mathcal{M}}$, too, compared to steady-state intervals. As a consequence, attenuating the disturbances in the former intervals is more imperative, especially in the constant velocity part, for which scanning takes place in the lithography process. Indeed, it can be directly observed in Figure 5 that the GP-based feedforward augmentation leads to significant improvement both in position tracking performance during steady-state, as well as position tracking performance during motion.

Finally, in order to quantify the resulting improvement with the GP-based feedforward augmentation, the ℓ_2 norm, scaled by the square root of the number of data points, and the ℓ_∞ norms of the $\hat{z}_{\mathcal{T}}^{\mathcal{M}}$ error are computed. In Table 3 the constant velocity results are presented, while in Table 2 the results for the whole trajectory are presented. Based on the results shown in both tables, it is asserted that the improvement is approximately 50%, thus highlighting the effectiveness of the proposed approach towards achieving higher positioning accuracy.

Table 2: Experimental results of the feedforward augmentation.

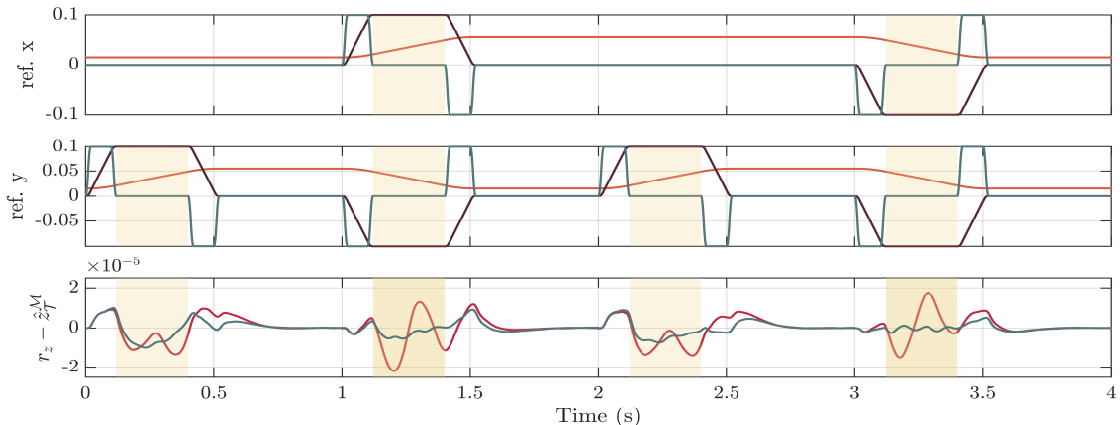|	No compensation	Kernel-based FF	Relative reduction
$\ell_2/\sqrt{N}$	$6.4 \cdot 10^{-6}$	$3.39 \cdot 10^{-6}$	47.03%
ℓ_∞	$2.19 \cdot 10^{-5}$	$9.97 \cdot 10^{-6}$	54.47%

Figure 5: From top to bottom: Reference trajectory in terms of position $[m]$ (━), velocity $[\frac{m}{s}]$ (━) and scaled acceleration $[\frac{m}{s^2}]$ by a factor of 10 (━) for $\hat{x}_{\mathcal{T}}^{\mathcal{M}}$. Reference position $[m]$ (━), velocity $[\frac{m}{s}]$ (━) and scaled acceleration $[\frac{m}{s^2}]$ by a factor of 10 (━) for $\hat{y}_{\mathcal{T}}^{\mathcal{M}}$. Error in z-direction $[m]$ with (━) and without (━) kernel-based feedforward compensation for static disturbance rejection. The yellow areas represent the time intervals for which the velocity is constant.

Table 3: Experimental results of the feedforward augmentation during constant velocity.

	No compensation	Kernel-based FF	Relative reduction
$\ell_2/\sqrt{N}$	$1.02 \cdot 10^{-5}$	$4.38 \cdot 10^{-6}$	57.06%
ℓ_∞	$2.19 \cdot 10^{-5}$	$9.97 \cdot 10^{-6}$	54.47%

5. Conclusions

This paper presents a kernel based modelling approach that is able to successfully capture the steady state behaviour of residual dynamics for moving-magnet planar actuator systems as a function of the generalized coordinates associated to the coil - metrology and the translator - metrology reference frames. Additionally, capturing the static behaviour of residual dynamics using the GP framework allows for augmentation of the feedforward controller, such that unforeseen static effects of residual dynamics are timely attenuated for. Consequently, position tracking error in $z_{\mathcal{T}}^{\mathcal{M}}$ direction is reduced by more than $50\,\%$ during application.

Finally, in order to further improve position tracking performance of a planar motor stage, the kernel based modelling strategy can be extended, such that the full dynamic residual model is considered for feedforward compensation.

References

H. Butler. Position control in lithographic equipment. *IEEE Control Systems Magazine*, 31(5): 28–47, 2011.

H.S. Cho, C. H. Im, and H. K. Jung. Magnetic field analysis of 2-d permanent magnet array for planar motor. *IEEE Transactions on Magnetics*, 37(5):3762–3766, Sep. 2001. ISSN 0018-9464. doi: 10.1109/20.952708.

G. M. Clayton, S. Tien, K. K. Leang, Q. Zou, and S. Devasia. A review of feedforward control approaches in nanopositioning for high-speed spm. *Journal of dynamic systems, measurement, and control*, 131(6):061101, 2009.

J. C. Compter. Electro-dynamic planar motor. *Precision Engineering*, 28(2):171–180, 2004.

C. Custers. *Overactuated magnetically levitated systems: flexible modes and eddy current phenomena in high-precision bearingless planar motors*. PhD thesis, Eindhoven University of Technology, 2019.

J. de Boeij, E. Lomonova, and J. Duarte. Contactless planar actuator with manipulator: A motion system without cables and physical contact between the mover and the fixed world. *IEEE Transactions on Industry Applications*, 45(6):1930–1938, 2009.

F. M. Dekking, C. Kraaikamp, H. P. Lopuhaä, and L. E. Meester. *A Modern Introduction to Probability and Statistics: Understanding why and how*. Springer Science & Business Media, 2005.

Martin Goubej, Sven Meeusen, Noud Mooren, and Tom Oomen. Iterative learning control in high-performance motion systems: from theory to implementation. pages 851–856, 09 2019. doi: 10.1109/ETFA.2019.8868996.

D. Jeltsema and J.M.A. Scherpen. Multidomain modeling of nonlinear networks and systems. *IEEE Control Systems*, 29(4):28–59, 2009.

P. Lambrechts, M. Boerlage, and M. Steinbuch. Trajectory planning and feedforward design for electromechanical motion systems. *Control Engineering Practice*, 13(2):145–157, 2005.

D. J. C. MacKay. Introduction to Gaussian processes. *NATO ASI Series F Computer and Systems Sciences*, 168:133–166, 1998.

Noud Mooren, Gert Witvoet, and Tom Oomen. Gaussian process repetitive control for suppressing spatial disturbances, 2020.

R. M. Murray, Z. Li, and S. S. Sastry. *A mathematical introduction to robotic manipulation*. CRC press, 1994.

R.M. Neal. *Bayesian learning for neural networks*, volume 118. Springer Science & Business Media, 2012.

Ioannis Proimadis. *Nanometer-accurate motion control of moving-magnet planar motors*. PhD thesis, Department of Electrical Engineering, 2020.

C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. MIT Press, 2005.

J. M. M. Rovers. *Multiphysical modeling of high-precision electromechanical devices*. PhD thesis, Eindhoven University of Technology, 2013.

J. M. M. Rovers, J. W. Jansen, J. C. Compter, and E. A. Lomonova. Analysis method of the dynamic force and torque distribution in the magnet array of a commutated magnetically levitated planar actuator. *IEEE Transactions on Industrial Electronics*, 59(5):2157–2166, 2012.

M. Steinbuch, R.J.E. Merry, M.L.G. Boerlage, M.J.C. Ronde, and M.J.G. Molengraft, van de. *Advanced Motion Control Design*, pages 27–1/25. CRC Press, 2010. ISBN 978-1-4200-7360-7. doi: 10.1201/b10382-35.

M. van de Wal, G. van Baars, F. Sperling, and O. Bosgra. Multivariable $\mathcal{H}_\infty/\mu$ feedback control design for high-precision wafer stage motion. *Control Engineering Practice*, 10(7):739–755, 2002.

Grace Wahba. *Spline models for observational data*, volume 59. Siam, 1990.