Model Parameter Identification for Feed Drive Dynamics using Kernel-Based Methods

by

James Duncan McPherson
B.A.Sc., University of Waterloo, 2021

A Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

MASTER OF APPLIED SCIENCE

in the Department of Mechanical Engineering

Model Parameter Identification for Feed Drive Dynamics using Kernel-Based Methods

by

James Duncan McPherson
B.A.Sc., University of Waterloo, 2021

Supervisory Committee

Dr. Keivan Ahmadi., Supervisor
(Department of Mechanical Engineering, University of Victoria, Canada)

Dr. Daniela Constantinescu., Departmental Member
(Department of Mechanical Engineering, University of Victoria, Canada)

ABSTRACT

This thesis presents an application of kernel-based methods for identification of the feed drive's dynamics model parameters and prediction of the disturbances affecting feed drives during operation from the cutting forces. To this purpose, the Partially Linear - Least Squares Support Vector Machine (PL-LSSVM) and Kernel Recursive Least Squares - Tracker (KRLS-T) algorithms were utilised for batch identification and online prediction of disturbances. Experimental case studies were performed with two ball screw feed drives under simulated and real cutting conditions to verify the identification and in-operation prediction of cutting forces.

Contents

List of Tables

List of Figures

ACKNOWLEDGEMENTS

I would like to thank:

My Mom and Dad, for supporting me through the entire process.

Dr. Ahmadi, for mentoring, support, encouragement, and extreme patience.

My fellow members of the DDM lab, for their friendship and support through this trial.

NSERC, for funding my research.

DEDICATION

To my goal, Finden.

Chapter 1

Introduction

1.1 Motivation

Feed drives are an essential part of Computer Numerical Controlled (CNC) machine tools. They are responsible for the transmission of movement from an electric motor to a work table or cutting tool. Feed drives consist of electrical and intermediate mechanical parts, coupled to each other to perform specific motion commands which move the workpiece or cutting tool to a desired position.

Different types of disturbances, such as forces from friction and cutting of material, can be present during the operation of the feed drives. It is important to be able to identify the nature of such disturbance forces in advance and use various pre-compensation techniques, i.e. feedforward control, to cancel them in a stable manner and minimize the tracking error [1]–[5]. Since these methods are dependent on the exact transfer function of the system, any mismatch between the actual and identified dynamic parameters of the feed drive leads to drastic degradation of the performance of the feedforward compensator [6], [7]. Therefore, precise identification of model parameters is critical to the optimal design of controllers for maximizing machine performance [8]–[10].

To this aim, various experimental methods have been developed for accurately identifying the parameters of linear feed drive's dynamics model with the use of experimental data. These methods can identify the parameters of the dynamics but require the measurements to be conducted during machine downtime to remove biasing effects, in particular cutting force disturbances, from the data [5], [11], [12]. The resulting model is usually not an accurate representation of the feed drive's dynamics during machining due to the absence of cutting forces. Furthermore, offline model parameters identification method are not suitable for continuously calibrating the

feed drives' digital twins, which are a novel concept increasingly being adopted for process monitoring and optimization. These limitations motivate the development of new methods for the online identification of the feed drive's dynamic model based on experimental measurements collected during normal machining operations.

1.2 Research Objectives

The objectives of this thesis are to:

1. Establish an algorithm that can identify the linear dynamics of feed drives during online operation of CNC machine tools.
2. Identify the disturbances in the feed drives caused by milling forces using the identified dynamics model and internal measurements from the controller.

To accomplish this, we will formulate the rigid body dynamics of a feed drive as a Partially Linear - Autoregressive model with Exogenous inputs (PL-ARX), where the linear parameters to be identified are derived from the physics of the feed drive's dynamics. The nonlinear milling force disturbances will be predicted using a locally periodic kernel simultaneously from in-process controller signals.

This thesis presents experimental results which will show, with the proper selection of hyperparameters, the linear feed drive parameters and the nonlinear cutting force disturbances can be identified.

1.3 Thesis Organization

This thesis will follow the structure as described below:

1. Chapter 2 reviews the existing literature regarding the theory of feed drive dynamics, and parametric identification of feed drive's dynamic models, including kernel-based model identification methods.
2. Chapter 3 presents the experimental case studies used to evaluate the kernel-based method.
3. Chapter 4 presents the detailed analysis of the results of this work and discusses the effectiveness of the kernel-based methods in identification and prediction.
4. Chapter 5 summarizes the significant contributions of this thesis and identifies the opportunities for future work.

Chapter 2

Literature Review

This chapter will review the relevant literature. In particular, it covers the existing models of ball screw feed drive dynamics, the identification of those models' parameters, and the application of kernel-based methods in improving those identification techniques.

2.1 Modeling of Feed Drive Dynamics

Industrial feed drives mainly consist of two different types, ball screw and linear feed drives. As seen in Fig. 2.1, a ball screw feed drive consists of a table which is connected to an electric motor through the coupling of a ball screw and nut. Linear feed drives replace the motor and ball screw with a series of magnets and wire winding's which moves the table in the same fashion the magnets and wire winding's inside the motor power its rotary motion. For this thesis we will be considering ball screw feed drives in our model.

Modeling feed drive dynamics, i.e. its velocity response to the input motor current, is essential for designing high-performance CNC. While various experimental methods exist to identify the feed drive's model parameters, their complexities depend highly on the model structure. Some models with high-order complex dynamics produce realistic simulations, but they demand sophisticated first-principle modeling combined with precise experimental procedures for identification [13]–[16]. On the other hand, simpler models, such as the one created by Erkorkmaz and Altintas, give the opportunity for rapid in-situ identification that can be used for real-time monitoring and process optimization [5], [6], [10].

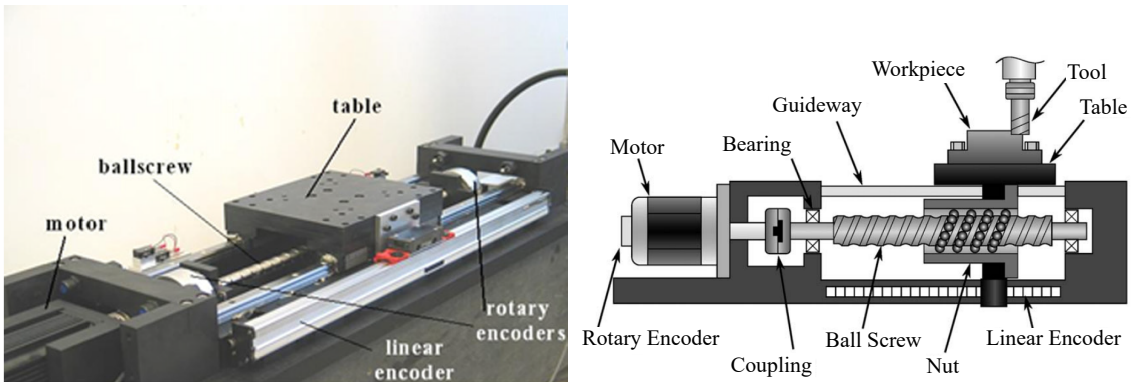

Figure 2.1: A Single-Axis Ball Screw Feed Drive [17]. Schematic of Components of a Typical Ball Screw Feed Drive [18].

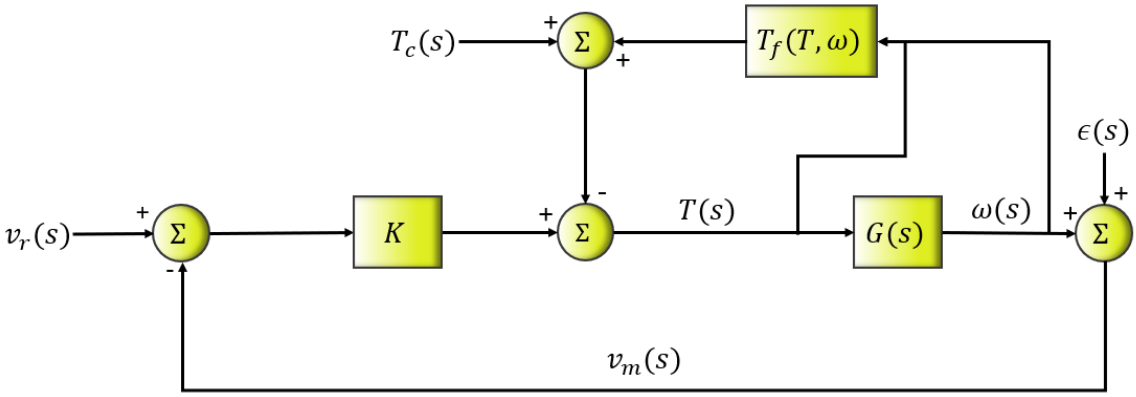

Figure 2.2: Velocity control loop of a typical feed drive. The variables of the control loop are as follows; v_r is the reference velocity input, K is the gain of a linear controller, T the actuation torque, G transfer of the feed drive dynamics, ω velocity of the feed drive, v_m measured velocity, ε measurement noise, T_c disturbance torque from cutting forces, and T_f disturbance torque from friction.

The velocity control loop of a typical feed drive, is shown in Figure 2.2. The reference velocity, v_r, is summed with the negative feedback from the output velocity, v_m. This is done digitally within the CNC controller and is used as the input for the control logic. The resulting velocity error is used to control the motor torque, this relationship is linear and is represented by the constant, K. The current control loop and back electromagnetic force of the motor is neglected. The motor torque is disturbed by the torque generated by a combination of frictional, T_f, and cutting forces, T_c. Measurement noise is assumed to be white and denoted by ε.

The resulting torque, the actuation torque T, then acts upon the feed drive which is represented by the transfer function, $G(s) = \frac{1}{Js+B}$, as per Altintas [17]. The inertia, $(kg.m^2)$, and viscous damping $(kg.m^2/s)$ parameters of the ball-screw feed drive are represented as J, and B respectively. This transfer function is a common choice for representing ball-screw and linear feed drives [11], [18].

Assuming Zero-Order-Hold discretization with $k = 1, ..., N$ time steps at T_s sampling interval the transfer function can be converted into the discrete domain with a gain, K_d, and a pole, p_d, as shown in Eq. 2.1. This creates the following discrete function, Eq. 2.2, for predicting the velocity output of the feed drive, $v_m(k+1)$.

$$\begin{aligned} K_v = \frac{K}{J}, \quad p = \frac{-B}{J} \\ K_d = \frac{K_v(p_d - 1)}{p}, \quad p_d = e^{pT_s} \end{aligned} \tag{2.1}$$

$$v_m(k+1) = (p_d - K_d)v_m(k) + K_d v_r(k) - \frac{K_d}{K_v}(T_c(k) + T_f(k)) + \varepsilon(k) \tag{2.2}$$

Accurately modelling friction complicates parameter identification because most popular friction models, like LuGre, are nonlinear in parameters and are identified by multiple controlled experiments and nonlinear optimization [19]–[23]. To avoid this, Altintas and Erkorkmaz developed an expression of the friction torque, T_f, based on the survey on the physics behind frictional forces by Armstrong et al. [5], [24]. The frictional forces are expressed as a piecewise linear function which is dependent on the axis velocity, ω, and the actuation motor, T. This piecewise representation of the friction is utilized to separate the Coulomb friction from Eq. 2.2 to create Eq. 2.3. The disturbance forces remaining after removing the friction are T_c, which represent the disturbance from the cutting of the material during machining.

$$v_m(k+1) = (p_d - K_d)v_m(k) + K_d v_r(k) - \frac{K_d}{K_v}(\bar{d}_p s^+ + \bar{d}_n s^-)T_f(k) - \frac{K_d}{K_v}T_c(k) + \varepsilon(k) \tag{2.3}$$

$$\text{where } s^{\pm} = \begin{cases} \pm 1, & \pm v_m(k) > 0 \\ 0, & \pm v_m(k) \leq 0 \end{cases} \tag{2.4}$$

This final model, Eq. 2.3, removes the nonlinear disturbance caused by the friction and cutting forces from the feed drive's dynamic model parameters. The inertia, J, and viscous damping,

B, are identified alongside the positive and negative velocity Coulomb friction, represented by $\bar{d}_p$ and $\bar{d}_n$ respectively. By applying identification techniques to this model the parameters of the feed drive dynamics can be identified.

2.2 Identification Methods

2.2.1 Least Squares

The equation of the feed drive's dynamics model, Eq. 2.3, can be rewritten in the form of a general Partially Linear - Autoregressive model with Exogenous inputs (PL-ARX), as seen by Erkorkmaz and Altintas [5]. The general PL-ARX model can be formulated in the following regression form:

$$y(k) = \boldsymbol{\beta}^T \boldsymbol{z}(k) + f(\boldsymbol{x}(k)) + \varepsilon(k) \tag{2.5}$$

where $y(k) \in \mathbb{R}$ is the system output at the discrete time step k. The vector $\boldsymbol{\beta} \in \mathbb{R}^m$ comprises the autoregressive and exogenous coefficients, and $\boldsymbol{z}(k) \in \mathbb{R}^m$ is the vector of the corresponding variables, i.e. the linear regressors. Similarly, $\boldsymbol{x}(k) \in \mathbb{R}^n$ is the vector of nonlinear regressors. The error term, $\varepsilon(k) \in \mathbb{R}$, is assumed to be i.i.d with zero mean random process. The exact formulation of the parameters of the feed drive's dynamics model is detailed in Eq. 2.6 and 2.7. Ω is a variable set to account for measurement noise of the velocity signal.

$$\boldsymbol{\beta} = \begin{bmatrix} K_d \\ (p_d - K_d) \\ \frac{K_d}{K_v}\bar{d}_p \\ \frac{K_d}{K_v}\bar{d}_n \end{bmatrix}; \quad \boldsymbol{z}_k = \begin{bmatrix} v_r(k) \\ v_m(k) \\ -\sigma_p(v_m(k)) \\ -\sigma_n(v_m(k)) \end{bmatrix}; \quad f(k) = -\frac{K_d}{K_v} T_c(k); \quad y(k) = v_m(k+1) \tag{2.6}$$

$$\begin{aligned} \sigma_p &= \frac{1}{2}\tau(v_m)\tau(1+\tau(v_m)) \\ \sigma_n &= \frac{-1}{2}\tau(v_m)\tau(1-\tau(v_m)) \\ \tau(v_m) &= \begin{cases} 1 & \text{if } v_m > \Omega \\ 0 & \text{if } |v_m| < \Omega \\ -1 & \text{if } v_m < \Omega \end{cases} \end{aligned} \tag{2.7}$$

When the disturbance, $f(k)$, is white noise, unbiased estimation of model parameters, $\boldsymbol{\beta}$, can be obtained from Least Squares (LS) approximation as shown in Eq. 2.8. To achieve this assumption identification tests are conducted with datasets collected from air-cutting trials [5], [11], [12], [25].

$$\begin{aligned} \boldsymbol{\beta} &= \arg\min_{\boldsymbol{\beta}} \ \|y - \boldsymbol{\beta}^T \boldsymbol{z}\|^2 \\ \boldsymbol{\beta} &= \left(\boldsymbol{z}^T \boldsymbol{z}\right)^{-1} \boldsymbol{z}^T y \end{aligned} \tag{2.8}$$

When the disturbance, $f(k)$, is not white the least squares approximation leads to biased parameters. The problem with air-cutting is the machine tool can not be used during the identification process which is not ideal in an industry focused on productivity. A better approach is to accurately identify the feed drive parameters during machining operation.

2.2.2 Partially Linear - Least Squares Support Vector Machines

Traditional system identification methods without the cutting forces present (i.e. during air-cutting) lead to the simple least square estimation of the model parameters $\boldsymbol{\beta}$. In this work, we utilized the Partially Linear - Least Squares Support Vector Machines method to estimate the system parameters under cutting conditions, i.e. when $f(k) \neq 0$, via the following linear regression [26]:

$$\begin{bmatrix} \boldsymbol{\Omega} + \gamma^{-1}\mathbf{I} & \mathbf{1} & \boldsymbol{Z} \\ \mathbf{1}^T & 0 & \mathbf{0}_{1\times p} \\ \boldsymbol{Z}^T & \mathbf{0}_{p\times 1} & \mathbf{0}_{p\times p} \end{bmatrix} \begin{bmatrix} \boldsymbol{\alpha} \\ c \\ \boldsymbol{\beta} \end{bmatrix} = \begin{bmatrix} \boldsymbol{y} \\ 0 \\ \mathbf{0}_{1\times p} \end{bmatrix} \tag{2.9}$$

where $\boldsymbol{Z} = [\boldsymbol{z}(1), ..., \boldsymbol{z}(N-1)]^T$, $\boldsymbol{y} = [v_m(2), ..., v_m(N)]^T$, $\mathbf{1}$ and $\mathbf{0}$ are vectors of ones and zeros of appropriate size, respectively, and $\boldsymbol{I}$ is identity matrix. The vector $\boldsymbol{\alpha} \in \mathbb{R}^N$ denotes Lagrangian multipliers, γ is the regularization coefficient, and $\boldsymbol{\Omega}$ is the kernel matrix:

$$\boldsymbol{\Omega}_{ij} = K(\boldsymbol{x}(i), \boldsymbol{x}(j)) \tag{2.10}$$

where $\boldsymbol{\Omega}_{ij}$ denoting the i, j entry of the kernel matrix and K is the kernel function defining the correlation between i and j data points [27]. The Moore-Aronszaj theorem states that for any positive semi-definite kernel function there is a unique Hilbert space that is characterized by its reproducing kernel, this is known as a Reproducing Kernel Hilbert Space (RKHS) [28]. To create a hypothesis space that can effectively represent the cutting force disturbances, periodic

kernels with varying amplitude are used:

$$K(i,j) = e^{-\frac{\|i-j\|^2}{2\sigma_1^2}} \times e^{-\frac{\sin\frac{\|i-j\|^2}{\psi}}{2\sigma_2^2}} \tag{2.11}$$

where σ_1, σ_2, and the regularization coefficient γ are the hyperparameters. This kernel spans all of the functions that are periodic at every ψ data point. Because the cutting force disturbances are periodic at every spindle rotation period (considering tool runout), this parameter is easily determined according to the known spindle speed. Using the $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and c parameters identified from the least squares estimation by Eq. 2.9, the disturbance term in the ARX model of Eq.2.5 can be estimated as follows:

$$f(k) = \sum_{i=1}^{N} \alpha_i K(k,i) + c \tag{2.12}$$

The application of the PL-LSSVM method leads to estimating the unbiased model parameters, $\boldsymbol{\beta}$, as well as the unknown periodic disturbance caused by the cutting forces. The estimated disturbance is only a scaled representation of the actual cutting forces and it does not include static forces because they are not observed in velocity output (i.e. $c \approx 0$).

2.2.3 Kernel Recursive Least Squares - Tracker

In the previous section, the disturbance for the entire measured length, $\boldsymbol{f}_k = [f(1)...f(N)]^T$, was estimated by the batch PL-LSSVM method in Eq.2.12. In this section, the goal is to predict $\boldsymbol{f}_k = [f(1)...f(k+1)]^T$, for any $k \leq N$ before a new observation $v_m(k+1)$ becomes available. This objective can be achieved by solving Eq.2.12 at each time step, but it requires recomputing all of the α_i for $i = 1..k+1$ each time, and the computation time grows cubically by k [29]. Instead, adopting the probabilistic approach in Gaussian Process Regression (GPR) [27], we define the disturbance as a Gaussian process in the following form:

$$v(k) = v_m(k) - \boldsymbol{\beta}^T \boldsymbol{z}(k) = f(k) + \varepsilon(k) \tag{2.13}$$

where $\boldsymbol{\beta}$ is deterministically known from the batch estimation. In a Bayesian framework, we assume the prior mean of $\boldsymbol{f}_k$ is zero and its covariance is equal to the kernel matrix, i.e. $p(\boldsymbol{f}_k) = \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}_k)$, where $\boldsymbol{\Omega}_k$ is the first k rows and columns of the total kernel matrix, $\sigma_0^2\boldsymbol{\Omega}$, and σ_0^2 is the disturbance variance. In standard Gaussian Process Regression (GPR), σ_0^2 is treated as a hyperparameter and learned by maximizing the marginal likelihood of all observations. In the recursive format of GPR, this parameter will be updated recursively as described in Eq.2.20. The

likelihood of observations $\boldsymbol{v}_k = [v(1), ..., v(k)]^T$ given $\boldsymbol{f}_k$ is assumed zero mean GP with diagonal covariance $\sigma^2 \boldsymbol{I}$. The posterior distribution $p(\boldsymbol{f}_k|\boldsymbol{v}_k) = \mathcal{N}(\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_k)$, is then obtained using standard batch GP updating or recursively, as follows [30]:

$$\begin{aligned}
\boldsymbol{\mu}_{k+1} &= \begin{bmatrix} \boldsymbol{\mu}_k \\ \hat{v}(k+1) \end{bmatrix} + \frac{v(k+1) - \hat{v}(k+1)}{\hat{\sigma}_v^2(k+1)} \begin{bmatrix} \mathbf{h}_{k+1} \\ \hat{\sigma}_f^2(k+1) \end{bmatrix} \\
\boldsymbol{\Sigma}_{k+1} &= \begin{bmatrix} \boldsymbol{\Sigma}_t & \mathbf{h}_{k+1} \\ \mathbf{h}_{k+1}^\top & \hat{\sigma}_f^2(k+1) \end{bmatrix} - \frac{1}{\hat{\sigma}_v^2}(k+1) \begin{bmatrix} \mathbf{h}_{k+1} \\ \hat{\sigma}_f^2(k+1) \end{bmatrix} \begin{bmatrix} \mathbf{h}_{k+1} \\ \hat{\sigma}_f^2(k+1) \end{bmatrix}^T
\end{aligned} \tag{2.14}$$

where $\mathbf{h}_{k+1} = \boldsymbol{\Sigma}_t \mathbf{q}_{k+1}$, $\mathbf{q}_{k+1} = \mathbf{Q}_k \boldsymbol{\sigma}_{k+1}$, $\mathbf{Q}_k = \boldsymbol{\Omega}_k^{-1}$, and $\boldsymbol{\sigma}_{k+1} = [K(1, k+1), K(2, k+1), ..., K(k, k+1)]^T$.

The inverse of the kernel matrix $\mathbf{Q}_k$ is also updated recursively using Eq. 2.15, where $\gamma_{k+1}^2 = K(k+1, k+1) - \boldsymbol{\sigma}_{k+1}^T \mathbf{Q}_t \boldsymbol{\sigma}_{k+1}$ is the projected uncertainty. This is a popular method to lower the time complexity to $O(k^2)$ by removing the calculation of the full kernel matrix inversion [29], [31]–[33].

$$\mathbf{Q}_{k+1} = \boldsymbol{\Omega}_{k+1}^{-1} = \begin{bmatrix} \mathbf{Q}_k & \mathbf{0} \\ \mathbf{0}^\top & 0 \end{bmatrix} + \frac{1}{\gamma_{k+1}^2} \begin{bmatrix} \mathbf{q}_{k+1} \\ -1 \end{bmatrix} \begin{bmatrix} \mathbf{q}_{k+1} \\ -1 \end{bmatrix}^\top \tag{2.15}$$

The recursion updates expressed in Eq. 2.14 and 2.15 are initialized in Eq. 2.16. This corresponds to inference of the proposed model according to a single data point and the inversion of the first kernel output as seen below.

$$\begin{aligned}
\mu_1 &= \frac{\boldsymbol{v}(1) K(x_1, x_1)}{K(x_1, x_1) + \sigma^2} \\
\Sigma_1 &= K(x_1, x_1) - \frac{K(x_1, x_1)^2}{K(x_1, x_1) + \sigma^2} \\
\mathbf{Q}_1 &= \frac{1}{K(x_1, x_1)}
\end{aligned} \tag{2.16}$$

In Eq. 2.14, $\hat{v}(k+1)$, and $\hat{\sigma}_v(k+1)$ denote the predictive mean and variance at $k+1$ before we know the new observation $v(k+1)$, i.e. $p(v_{k+1}|\boldsymbol{v}_k) = \mathcal{N}(\hat{v}(k+1), \hat{\sigma}_y(k+1))$. Similarly, $\hat{\sigma}_f^2$ is the predictive variance of disturbance, $f(k+1)$. These variables are updated each iteration as per Eq. 2.17.

$$\begin{aligned}\hat{v}(k+1) &= \mathbf{q}_{k+1}^T \boldsymbol{\mu}_k \\ \hat{\sigma}_f^2 &= \gamma_{k+1}^2 + \mathbf{q}_{k+1}^T \mathbf{h}_{k+1} \\ \hat{\sigma}_v(k+1) &= \sigma^2 + \hat{\sigma}_f(k+1)\end{aligned} \tag{2.17}$$

The predictive mean can also be reformulated as a weighted sum of the covariance between $k+1$ and all of the past time instances.

$$\hat{v}(k+1) = \mathbf{q}_{k+1}^T \boldsymbol{\mu}_t = \boldsymbol{\sigma}_{k+1}^T \boldsymbol{\alpha}_k \tag{2.18}$$

The kernel weights are $\boldsymbol{\alpha}_k = \mathbf{Q}_k \boldsymbol{\mu}_k$. In other words, the mean and variance of the posterior distribution, $\boldsymbol{\mu}_k$ and $\boldsymbol{\Sigma}_k$, form a *dictionary* of *basis* for the regression of the subsequent predictive mean. While all of the past data inform the estimation of future output of time-independent systems, in time-varying systems, like the machining process, the older information (or corresponding bases in the dictionary) undermines the accuracy of predictions. To avoid this, an appropriate forgetting strategy must be implemented at each iteration. We use the Back-to-the-Prior (B2P) forgetting strategy, Eq. 2.19, where the posterior mean and variance are diluted after each update by mixing them with the prior [30].

$$\begin{aligned}\boldsymbol{\Sigma}_k &\leftarrow \lambda \boldsymbol{\Sigma}_k + (1-\lambda)\boldsymbol{\Omega}_k \\ \boldsymbol{\mu}_k &\leftarrow \sqrt{\lambda}\boldsymbol{\mu}_k\end{aligned} \tag{2.19}$$

λ is the forgetting factor that is s.t. $\lambda \in (0, 1]$, though $\lambda \in [0.95, 1]$ is a more sensible range [30]. This is because a smaller forgetting factor forces the dictionary content towards the prior, eliminating the information gained from the observations.

The maximum likelihood estimation of the disturbance variance, σ_0^2, is also obtained recursively by incorporating the forgetting factor, as follows:

$$\begin{aligned}\sigma_0^2(t+1) &= \frac{N(t+1)}{D(t+1)}; \\ N(t+1) &= N(t) + \lambda \frac{v(t+1) - \hat{v}(t+1)}{\hat{\sigma}_v(t+1)}; N(1) = \frac{v(1)^2}{1+\sigma_n^2} \\ D(t+1) &= D(t) + \lambda; D(1) = 1\end{aligned} \tag{2.20}$$

As described in Eq. 2.14, one new basis is added to the dictionary at each iteration, leading to quadratically increasing computational cost. To restrict the computation cost to a constant value, the dictionary must be *pruned* at each iteration to eliminate the least informative basis, in

the literature this is also known as sparsification [29], [32], [34], [35]. Various methods can be used for pruning the dictionary; we use the fixed budget method presented by Van Varenbergh [30]. Here the dictionary size is kept constant at M by removing the basis with the minimum $\left(\frac{\left[\mathbf{Q}_{k+1}\boldsymbol{\mu}_{k+1}\right]_i}{[\mathbf{Q}]_{i,i}}\right)^2$, with $[.]_{i,j}$ indicating the entry at the i^{th} row and j^{th} column of the matrix. The candidate basis is removed from the dictionary by simply eliminating the corresponding row and column from posterior mean and variance. Removal of the basis from the inverted kernel matrix, $\mathbf{Q}_{k+1}$, is more involved but follows the same logic. The pruning is conducted by Eq. 2.21 and 2.22 as follows:

$$\begin{aligned} \boldsymbol{\mu}_{k+1} &\leftarrow \left[\boldsymbol{\mu}_{k+1}\right]_{-i} \\ \boldsymbol{\Sigma}_{k+1} &\leftarrow \left[\boldsymbol{\Sigma}_{k+1}\right]_{-i,-i} \end{aligned} \tag{2.21}$$

$$\mathbf{Q}_{k+1} \leftarrow \left[\mathbf{Q}_{k+1}\right]_{-i,-i} - \frac{\left[\mathbf{Q}_{k+1}\right]_{-i,i}\left[\mathbf{Q}_{k+1}\right]_{-i,i}^{\top}}{\left[\mathbf{Q}_{k+1}\right]_{i,i}} \tag{2.22}$$

where $[.]_{-i,-j}$ indicates removing the i^{th} row and j^{th} column of the matrix, or vector if no j is specified. The entire RKLS algorithm with a fixed budget dictionary and B2P prior is described in Alg. 1.

Algorithm 1 Kernel Recursive Least Squares Tracker (KRLS-T) [30]

Observe the initial data $\boldsymbol{x}(1), \boldsymbol{v}(1)$
Set hyperparameters for the kernel, regularization γ, forgetting factor λ, and dictionary size M
Initialize $\mu_1, \Sigma_1, \mathbf{Q}_1$ using Eq. 2.16
Add $\boldsymbol{x}(1)$ to the dictionary; $\boldsymbol{x}(k)_D = [\boldsymbol{x}(1)]$
for $k = 2, 3, ..., N$ **do**
    Compute μ_k, Σ_k using Eq. 2.19
    Observe new data $\boldsymbol{x}(k), \boldsymbol{v}(k)$
    Compute $\boldsymbol{k}_{k+1}$, between $\boldsymbol{x}(k)$ and every basis in the dictionary, $\boldsymbol{x}(k)_D$
    Compute k_{k+1}, between $\boldsymbol{x}(k)$ and $\boldsymbol{x}(k)$
    Compute $\boldsymbol{q}_{k+1},\ \hat{v}(k),\ \mathbf{h}_{k+1}$
    Update $\mu_{k+1}, \Sigma_{k+1}, \mathbf{Q}_{k+1}$ as per Eq. 2.14 and 2.15
    Add $\boldsymbol{x}(k)$ to the dictionary; $\boldsymbol{x}(k+1)_D = [\boldsymbol{x}(k)_D; \boldsymbol{x}(k)]$
    if $S > M$ (where S is the size of $\boldsymbol{x}(k+1)_D$) $||$ $\gamma^2_{k+1} < \epsilon$ (for some $\epsilon > 0$ close to machine precision) **then**
        if $\gamma^2_{k+1} < \epsilon$ **then**
            Index, i, is set to; $i = S$
        else
            Determine index, i; s.t. $\min \left(\frac{\left[\mathbf{Q}_{k+1}\boldsymbol{\mu}_{k+1}\right]_i}{[\mathbf{Q}]_{i,i}} \right)^2$
        end if
        Remove the i index from $\mu_{k+1}, \Sigma_{k+1}, Q_{k+1}$ using Eq. 2.21 and 2.22.
        Remove the i basis from the dictionary; $\boldsymbol{x}(k)_D$
    end if
end for

2.3 Summary

Previous research has shown that the design of high-performance controllers for CNC machines is dependent on the proper identification of feed drive's dynamic model parameters. The general equation of the feed drive's motion can be considered as a PL-ARX system. The linear portion describes the rigid body motion, while the nonlinear portion describes the disturbance from cutting of the material during machining processes. In air-cutting the identification of the linear system returns unbiased estimations as the milling disturbances are not present. When the milling disturbances are present, kernel-based identification techniques are able to return unbiased estimations of the feed drive's dynamic model parameters and the milling disturbances. For online operation, where time complexity is a large concern, recursive kernel-based methods allow for similar estimation of the milling disturbances.

Chapter 3

Experimental Method

This chapter will cover the experimental methodology of the two case studies that will be used in this thesis to validate the kernel-based methods. The first case study consists of a linear feed drive with a shaker mounted to produce forces representing cutting operations. This case study is a good test of the algorithm's ability to compensate for periodic disturbances on a feed drive. The second case study consists of slot milling 7050-T7451 Aluminium on a CNC mill. This is a more realistic case study as the end goal is to develop solutions for industrial CNC milling, which has more inherent complexities than the shaker can simulate.

3.1 Case Study #1: Shaker Disturbance on Linear Feed Drive

The purpose of this case study is to validate the ability of the kernel-based methods to remove bias caused by the periodic forces during the identification of the parameters of a feed drive's dynamic model. This was conducted by Dr. Keivan Ahmadi at the Technical University of Munich (TUM) utilizing a linear feed drive with a shaker mounted on top of it, shown in Figure 3.1, to generate the periodic disturbance.

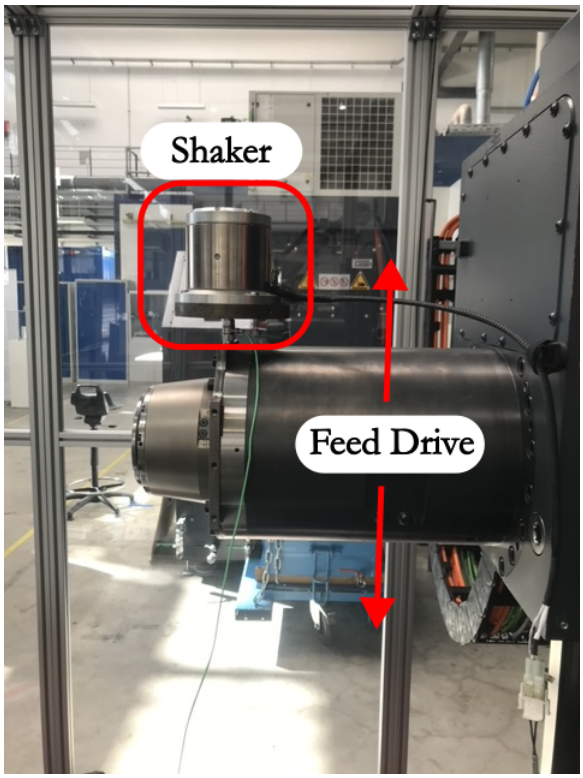

Figure 3.1: The setup for case study #1. A shaker is placed on the end of a linear feed drive which moves in the z axis described by the arrow

The shaker was programmed to disturb the feed drive and bias the identification of the linear components in the feed drive's dynamic model. All disturbances were conducted at a frequency of $55\ Hz$. This was chosen with regard to the Nyquist frequency of the measurement equipment, $167\ Hz$, to observe both the prime and secondary harmonics of the disturbance. Three different trials were conducted with this setup. Trial #1 was conducted as a control trial with a set trajectory designed for identification. Disturbances were then applied to the tests to bias the identification results. These disturbances were a constant amplitude, trial #2, and sweeping amplitude waveform, trial #3. The period of the disturbance for the latter trials was $55\ Hz$ which was used to determine the ψ hyperparameter of the kernel.

Each of these trials had the same velocity trajectory, seen in Figure 3.2. This trajectory was selected based on past identification literature, in particular [5]. This allowed for all linear parameters as described in section 2.1 to be identified.

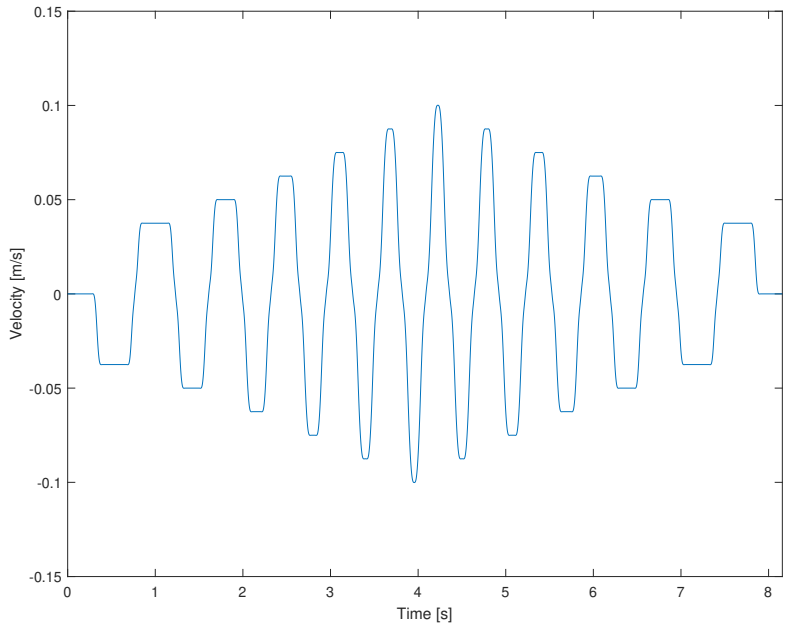

Figure 3.2: The reference velocity trajectory used in all case study #1 trials.

3.2 Case Study #2: Aluminium Slot Milling on CNC

The purpose of this experiment is to validate the capability of the kernel-based methods on milling forces. This is important as the main application of this work is CNC manufacturing. This work was performed at the Manufacturing Automation Laboratory (MAL) and was conducted by a member of MAL, Nima Dabiri. Slot milling trials were conducted on 7050-T7451 aluminium using a 2-fluted, 16 mm diameter flat endmill. The exact setup is shown in Figure 3.3.

The setup in Figure 3.3 shows the aluminium workpiece mounted in the CNC on top of a dynamometer (Kistler 9255B Multicomponent Dynamometer). This is a commonly used sensor to directly measure the milling forces. The axis reference velocity and encoder signals were collected using in-house developed software, INTELCUT. This application communicates with the Heidenhain CNC controller via TNC Ethernet connection to record controller's internal signals at 10 kHz sampling frequency. We down-sampled the recorded signals to 333 Hz, consistent with the control loop's time intervals.

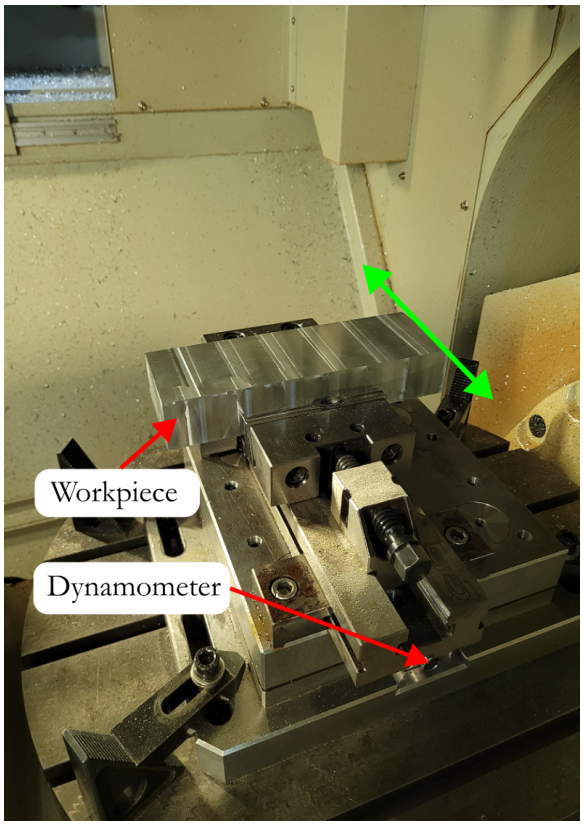

Figure 3.3: The setup for case study #2. The aluminium workpiece, dynamometer, and direction of motion are labeled within the CNC mill. The direction of motion for the X axis is represented by the double-headed green arrow.

Down-sampling the data imposed a limit on the maximum spindle speed that could be selected for the trials due to the Nyquist frequency. As the selected tool had two teeth, the first tooth passing frequency would be at twice the frequency of the spindle. So if a spindle speed of 6000 rev/min was selected, the spindle frequency would be at 100 Hz and the first tooth passing frequency will be at 200 Hz. As the Nyquist frequency is 167 Hz, a first tooth passing frequency of 200 Hz could not be recorded in the down-sampled data. Therefore the spindle speed was selected as 3000 rev/min for all milling trials as both spindle rotation frequency, 50 Hz, and first tooth passing frequency, 100 Hz, would be able to be measured after down-sampling.

Four tests were conducted with this setup, and each trial's parameters are shown in Table 3.1. In Trial 1 the tool and workpiece were not engaged (i.e. air cutting) to serve as control trials; in Trials 2, 3, and 4, the tool and workpiece were engaged with a constant 1.5 mm depth generating substantial machining forces in the machine's X, Y, and Z directions. In all of the trials the feedrate was calculated using a spindle speed of 3000 rev/min, with the machine's X-axis moving

at the programmed feedrate shown in 3.4.

Trial #	Milling Forces [RPM]	Trajectory
1	Air-Cutting [0]	Increasing/Decreasing Velocity
2	Constant Spindle Speed [3000]	Increasing/Decreasing Velocity
3	Constant Spindle Speed [3000]	Increasing Velocity
4	Constant Spindle Speed [3000]	Decreasing Velocity

Table 3.1: The different combinations of experimental parameters for the four trials. Whether milling forces were present, and the trajectory used for the trial.

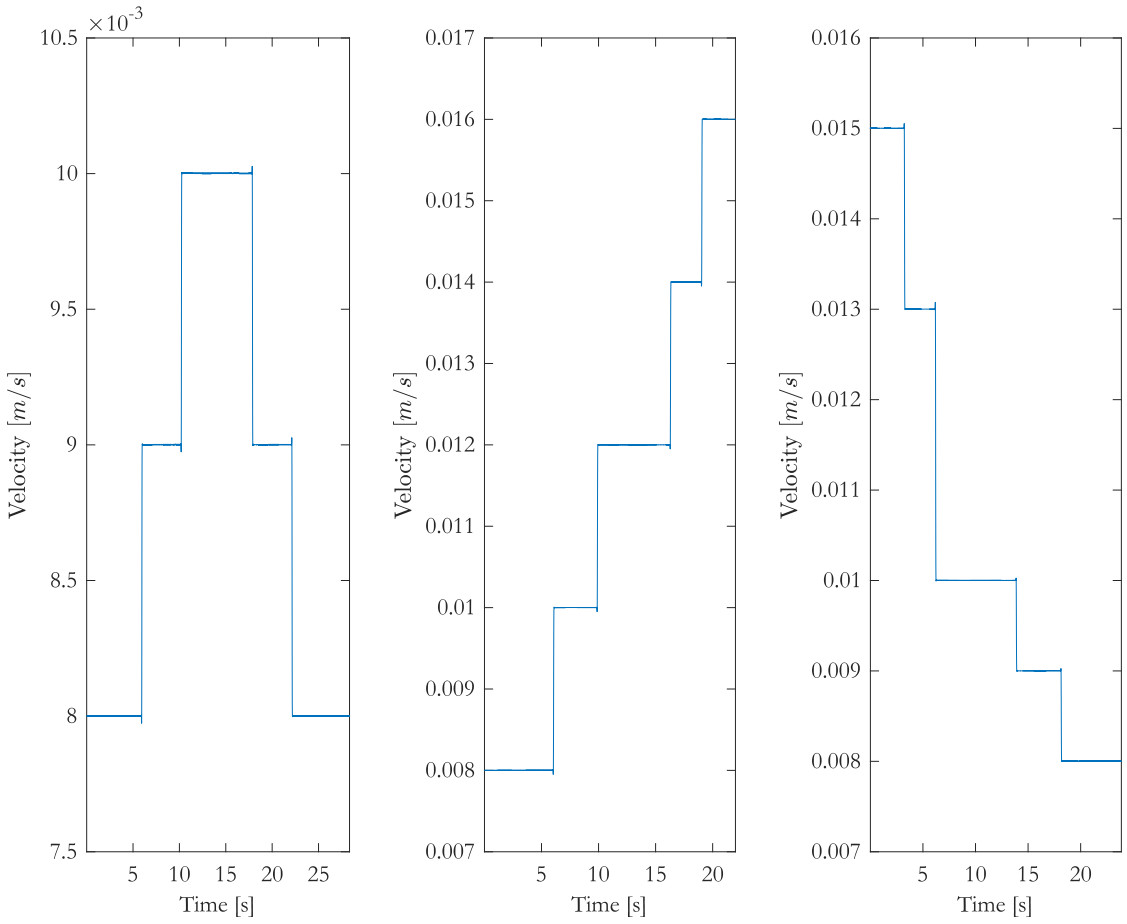

Figure 3.4: From left to right are the increasing/decreasing trajectory, the increasing trajectory, and the decreasing trajectory used in the case study with milling forces.

Trial 1 and 2 were conducted with the same trajectory for direct comparison of the bias in the parameter identification while trials 3 and 4 were conducted with different trajectories, as seen in Figure 3.4.

3.3 Summary

This chapter covered the case studies analyzed in this thesis and the identification using the PL-LSSVM method. The two experimental case studies presented in this chapter will allow for a thorough investigation of the PL-LSSVM and KRLS-T kernel-based methods. The first case study presents a typical identification trajectory with controlled nonlinear disturbance. The second case study builds on this by presenting a more typical trial in industry, where the tool path is not optimized for identification and the nonlinear disturbance is that of actual cutting forces. Finally, the ability of the PL-LSSVM method to identify the linear parameters of the feed drive dynamics model and the nonlinear disturbance acting on the feed drive was investigated with these case studies. The abilities of the PL-LSSVM identification method were found to be sufficient for the identification of both the linear and nonlinear aspects of the feed drive model.

Chapter 4

Analysis and Discussion

In this chapter, the kernel-based methods for batch identification of the feed drive's dynamic parameters and online prediction of the cutting force disturbances are applied to the two case studies described in section 3.1 and 3.2. In section 4.1 the PL-LSSVM method is used for the identification of the parameters which are validated by comparing the measured velocity output to the predictive output of a model built with the parameters. In section 4.2 the KRLS-T method is utilized for recursive identification of the cutting force disturbances on the feed drive. The linear components of the feed drive's dynamics model are compensated during the recursive operation using the parameters identified previoulsy with the PL-LSSVM method.

4.1 Partially Linear - Least Squares Support Vector Machines (PL-LSSVM)

4.1.1 Case Study #1: Shaker Test

The hyperparameters used for the PL-LSSVM method and the kernel function, Eq. 2.11, are shown in Table 4.1. All of these hyperparmeters were selected through hand tuning, except for the ψ hyperparameter. This was set as the period of the disturbance, which for the shaker disturbance used in the first case study was calculated to be 0.0182 seconds.

Hyperparameter	Value
σ_1	0.59
σ_2	0.42
γ	50.3
ψ	0.0182

Table 4.1: Hyperparameters utilized in PL-LSSVM of Case Study #1

The parameters of the feed drive's dynamics identified utilizing the LS and PL-LSSVM methods are shown in Table 4.2. The parameters identified from the trials containing the constant and sweeping disturbance, trials #2 and #3, are close to the parameters identified during the control trial, trial#1. The LS parameters have a larger variance due to the presence of the disturbance force. These results validate the ability of the PL-LSSVM method for unbiased identification of the feed drive's dynamic parameters. Figures 4.1, 4.2, and 4.3 show in greater detail the identification and prediction results of the PL-LSSVM method.

Trial #	#1		#2		#3	
Method	LS	PL-LSSVM	LS	PL-LSSVM	LS	PL-LSSVM
β_1	0.8338	0.8335	0.8317	0.8334	0.8327	0.8335
β_2	0.1632	0.1632	0.1651	0.1633	0.1646	0.1632
$\beta_3[1 \times 10^{-4}]$	-1.02	-1.40	-1.13	-0.906	-0.845	-0.922
$\beta_4[1 \times 10^{-4}]$	1.03	1.14	1.14	1.70	0.860	1.68

Table 4.2: Dynamic Model Parameters Identified with LS and PL-LSSVM for Case Study #1

The (d) panels illustrate the prediction error, $y - \hat{y}$, of the models created with parameters identified utilizing LS, and PL-LSSVM. Here the power spectral density (PSD) of the prediction error for both LS and PL-LSSVM are shown. For Figures 4.2, and 4.3 we can see the LS prediction error has a large amount of error at the disturbance frequency, $55\ Hz$, and its second harmonic, $100\ Hz$, compared to PL-LSSVM. This shows the PL-LSSVM's robust ability to model the disturbance frequency compared to LS.

Panels (a), (b), and (c) detail the fitness of the feed drive model using the identified linear parameters and trained kernel function versus the velocity and force measurements. Panel (a) shows the measured velocity of the feed drive overlaid with the predicted velocity of the model. As can be seen in all figures, the two velocities have a close fit with no visible differences in the plot's magnitude. To discern the fitness better, panel (b) shows the error between the measured

and predicted velocity. Across all figures, the error has a magnitude of approximately $10^{-4}[m/s]$, except Figure 4.2 where the error spikes when the shaker disturbance begins and ends abruptly. This shows that the selected kernel function is less accurate when predicting sharp changes in the disturbance.

Panel (c) details the output of the kernel portion of the model versus the measured forces from the shaker. Figure 4.1 has no disturbance so the kernel function attempts to fit any noise with the selected disturbance frequency. Figures 4.2, and 4.3 show the ability of the kernel to predict the disturbance compared to the measured force applied by the shaker. While the exact amplitude of the predicted force cannot be calculated, both figures show that the predicted disturbance matches the form of the measured force.

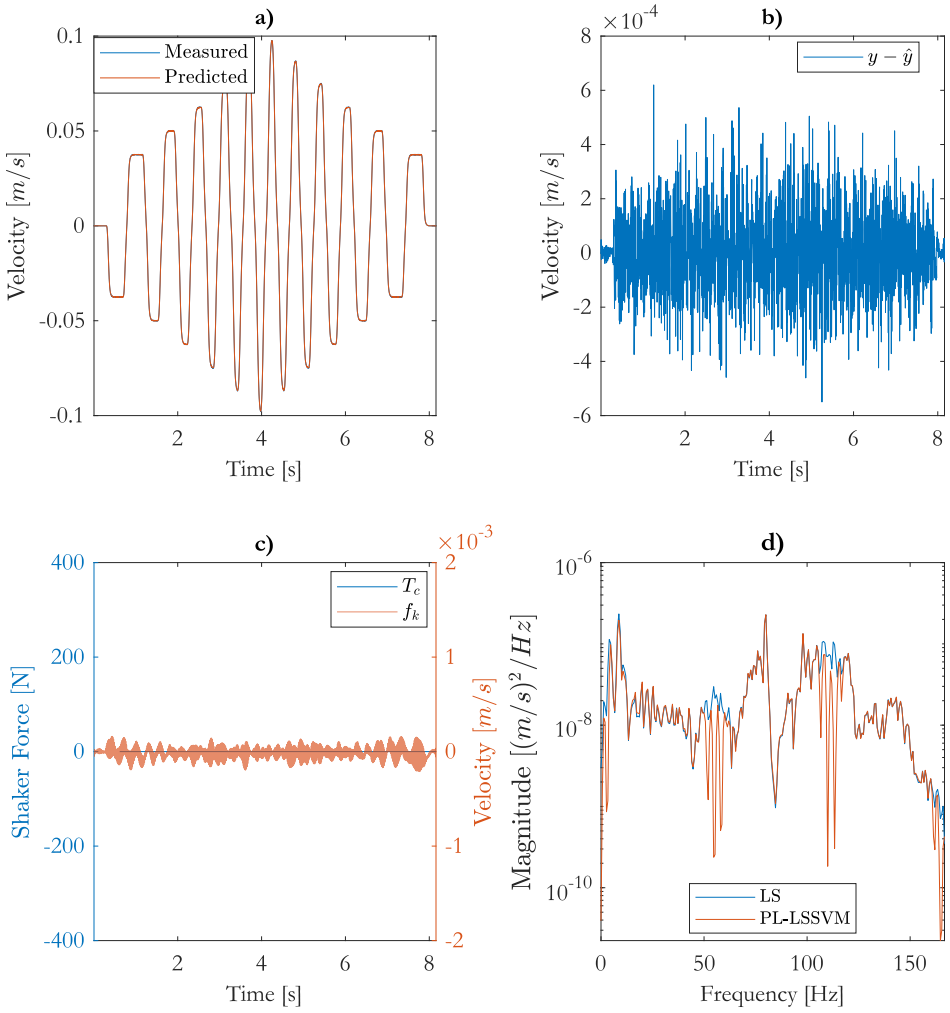

Figure 4.1: Results from the control trial of the shaker case study using PL-LSSVM for system identification and prediction. **a)** Measured output of the system vs. the prediction by the PL-LSSVM model. **b)** Error between the measured output and the PL-LSSVM model prediction. **c)** Output from the shaker vs. the identified nonlinear prediction. **d)** Power spectral density of the prediction error of Least Squares and PL-LSSVM.

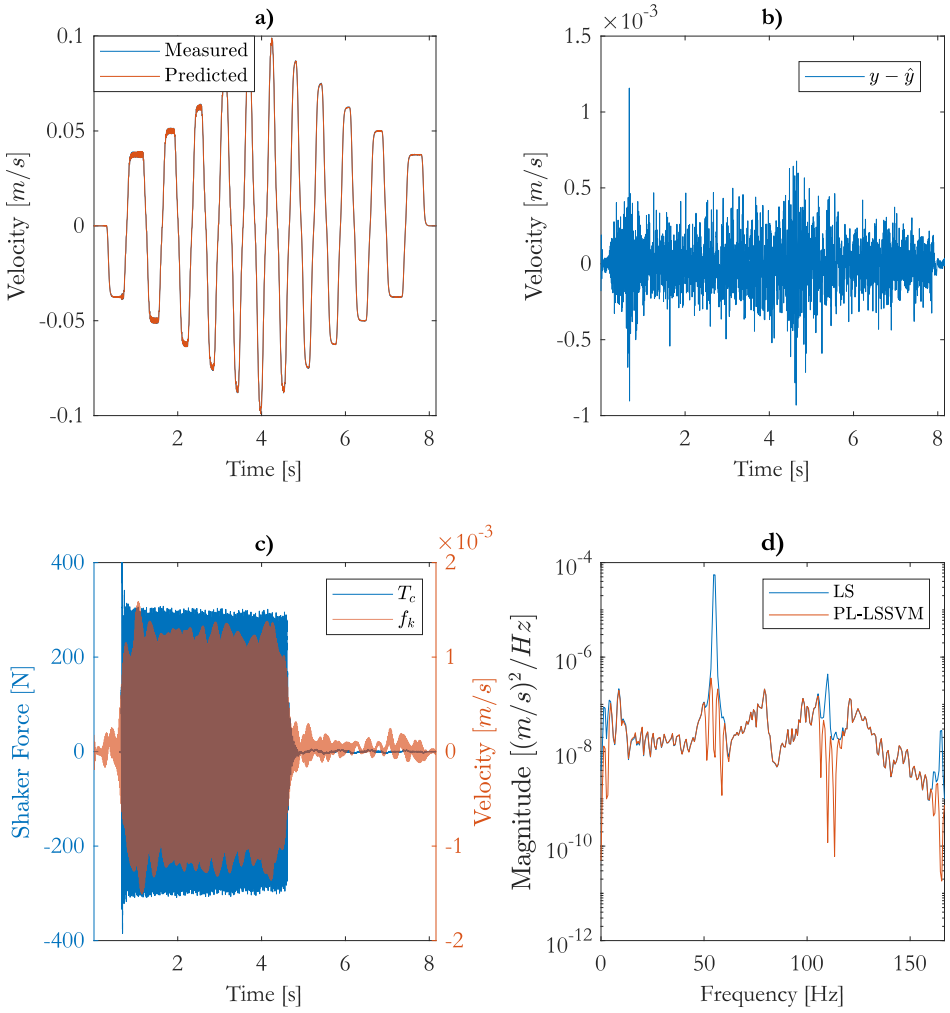

Figure 4.2: Results from the constant disturbance trial of the shaker case study using PL-LSSVM for system identification and prediction. **a)** Measured output of the system vs. the prediction by the PL-LSSVM model. **b)** Error between the measured output and the PL-LSSVM model prediction. **c)** Output from the shaker vs. the identified nonlinear prediction. **d)** Power spectral density of the prediction error of Least Squares and PL-LSSVM.

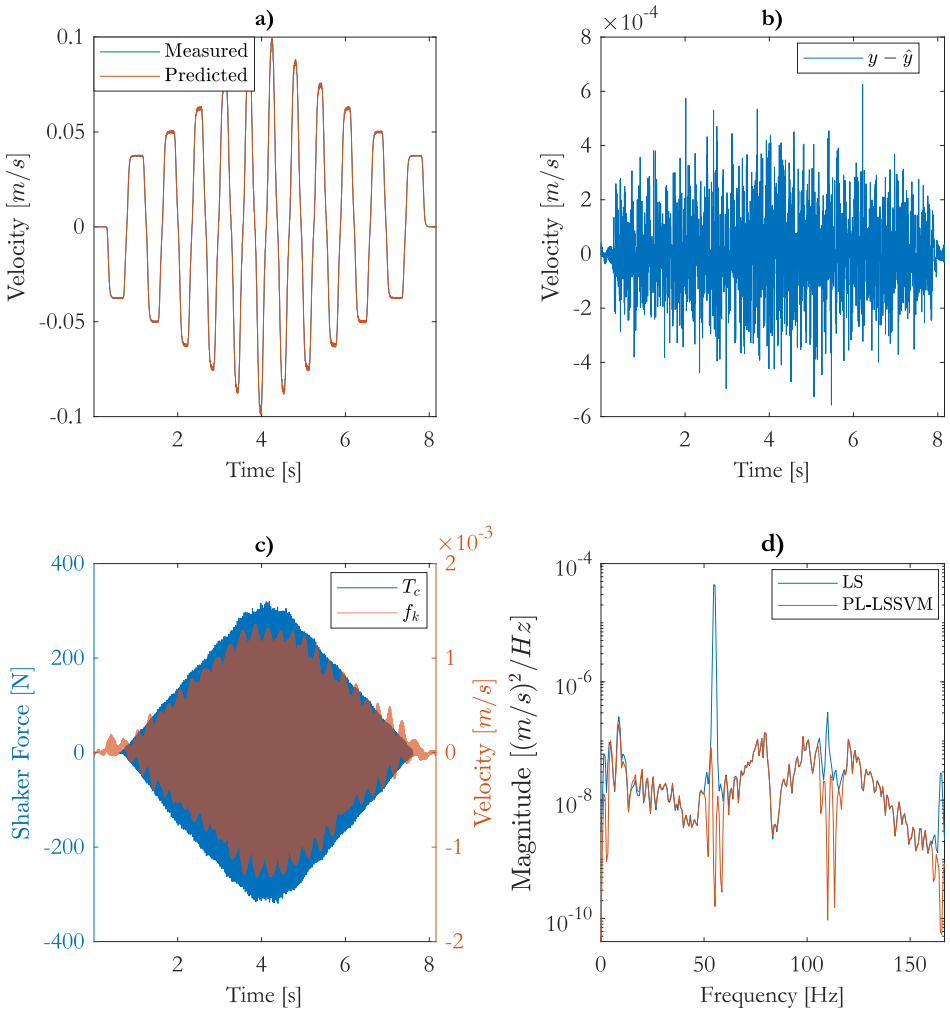

Figure 4.3: Results from the sweeping disturbance trial of the shaker case study using PL-LSSVM for system identification and prediction. **a)** Measured output of the system vs. the prediction by the PL-LSSVM model. **b)** Error between the measured output and the PL-LSSVM model prediction. **c)** Output from the shaker vs. the identified nonlinear prediction. **d)** Power spectral density of the prediction error of Least Squares and PL-LSSVM.

4.1.2 Case Study #2: Milling Test

The hyperparameters used for the periodic kernel and the PL-LSSVM method are shown in Table 4.1. As done with the first case study the σ_1, σ_2, and γ hyperparameters were selected through hand tuning and ψ was set as the disturbance period. In this case study it was calculated using the RPM of the spindle.

Hyperparameter	**Value**
σ_1	0.126
σ_2	0.42
γ	56
ψ	0.02

Table 4.3: Hyperparameters utilized in PL-LSSVM of Case Study #2

The parameters of the feed drive's dynamics that were identified with the LS and PL-LSSVM methods are shown in Table 4.4. For this case study the parameters of the Coulomb friction were not identified because the tested machining toolpaths do not include speed reversals. The Coulomb component of the friction force was negligible and therefore omitted from the equation of motion, leaving only two parameters in the linear portion of the model in Eq. 2.5. The two linear parameters of the model were identified using PL-LSSVM and LS method (neglecting disturbance). Additionally, the trajectories of the 3rd and 4th trial were not suitable for identification as they did not include both increasing and decreasing velocity profile for full excitation of the feed drive dynamics.

As with the previous results, the feed drive's dynamic parameters that are identified by the LS method are biased between the air-cutting and cutting trials. The parameters identified utilizing PL-LSSVM are more consistent though are not as unchanging as the results of the first case study. This is acceptable due to the bias by the frictional parameters not being identified. These results show that the PL-LSSVM can remove the nonlinear bias caused by milling forces from the identification of the feed drive's dynamic parameters.

Trial #	#1		#2	
Method	LS	PL-LSSVM	LS	PL-LSSVM
β_1	0.3473	0.3699	-0.2663	0.3722
β_2	0.6527	0.6292	1.2666	0.6278

Table 4.4: System Parameters Identified with LS and PL-LSSVM methods for Trials #1 and #2

Figures 4.4, 4.5, 4.6, and 4.7 show the ability of the PL-LSSVM method to identify the linear and nonlinear components of the feed drive's dynamics model. As with the previous results, panel (d) of Figures 4.5, 4.6, and 4.7 show the reduction in the disturbance frequency of $50\ Hz$ and the first tooth passing frequency of $100\ Hz$. This indicates that milling disturbances can be identified by the kernel. Additional reductions in magnitude can be seen at approximately $16.5\ Hz$ and $33\ Hz$. This correlates to $1/3$ and $2/3$ of the disturbance frequency. This is due to the interplay of the sampling frequency of the CNC controller, the bandwidth of the downsampling filter, and the frequency of disturbance used in the kernel function.

Panel (a) shows the measured velocity and the predicted velocity of the model. The results in Figures 4.4, 4.5, 4.6, and 4.7 are similar to the previous case study with regards to the overall prediction of the feed drive's velocity output.

The results of the velocity error, shown in panel (b), illustrate the error between the measured and predicted velocity. Across all figures the error has a magnitude of approximately $10^{-5}[m/s]$, through Figures 4.5, 4.6 and 4.7. These figures contain sudden jumps in the prediction error. This is caused by an error in the measurement software which failed to update the output at certain time steps, and as such does not indicate any problem with the PL-LSSVM method.

Panel (c) details the output of the kernel portion of the model versus the measured forces from the endmill. The constant component of the measured forces are removed for better comparison as the prediction of the velocity disturbance do not include this component. Figure 4.4 have no disturbance so the kernel attempts to fit to any noise with the selected disturbance frequency. Figures 4.5, 4.6, and 4.7 show the ability of the kernel to predict the disturbance compared to the measured force applied by the endmill. As with the first case study, the predicted disturbance matches the overall waveform of measured cutting forces, though is not scaled due to lacking the required parameters.

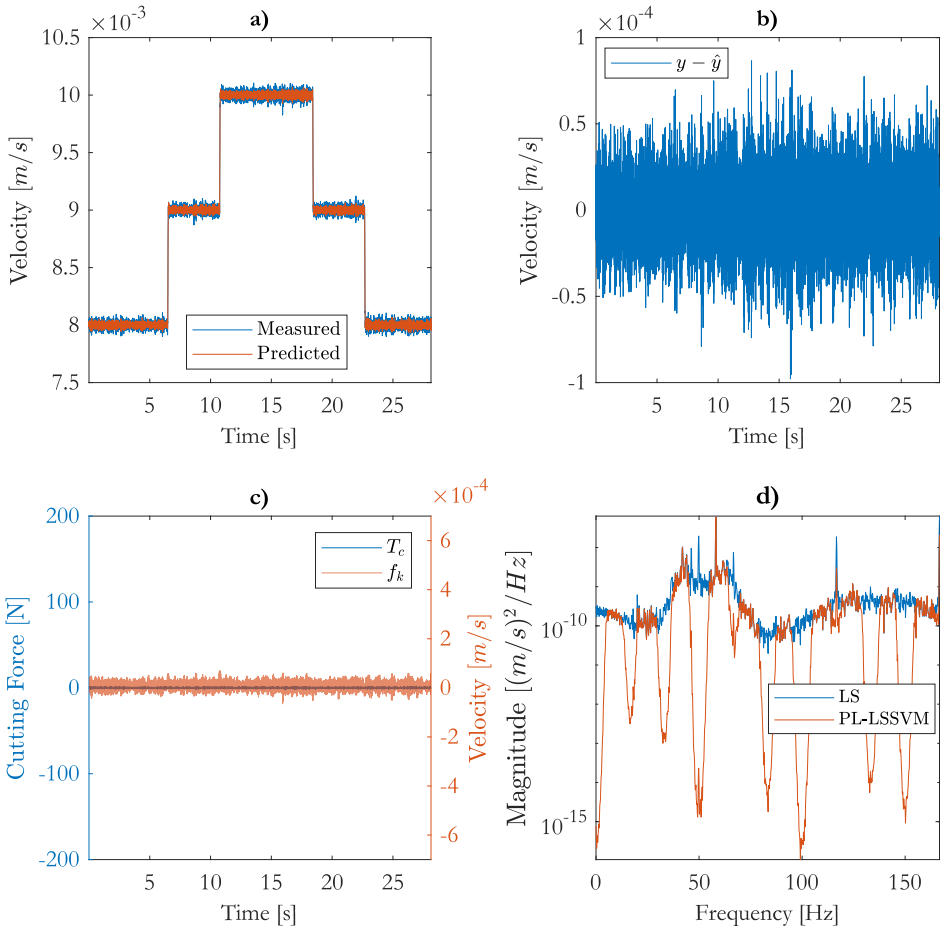

Figure 4.4: Results from the control trial of the milling case study using PL-LSSVM for system identification. **a)** Measured output of the system vs. the prediction by the PL-LSSVM model. **b)** Error between the measured output and the PL-LSSVM model prediction. **c)** Output from the endmill vs. the identified nonlinear prediction. **d)** Power spectral density of the prediction error of Least Squares and PL-LSSVM.

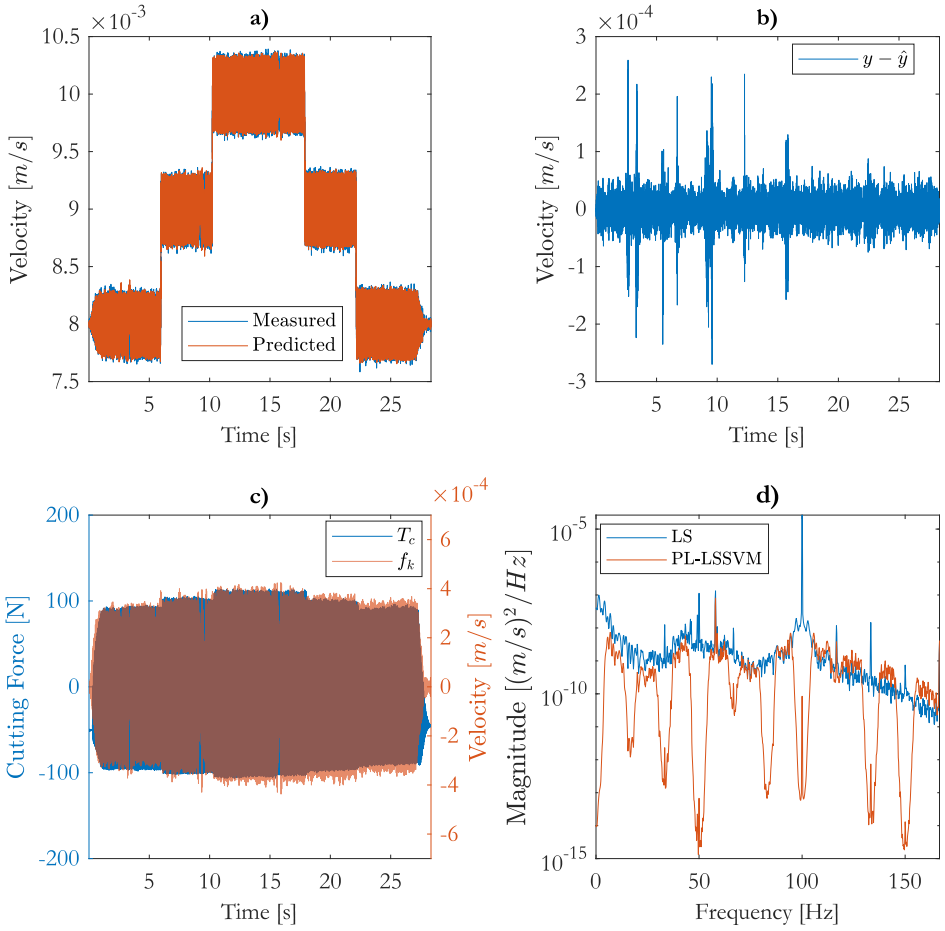

Figure 4.5: Results from the milling trial with increasing and decreasing velocity trajectory from the milling case study using PL-LSSVM for system identification. **a)** Measured output of the system vs. the prediction by the PL-LSSVM model. **b)** Error between the measured output and the PL-LSSVM model prediction. **c)** Output from the endmill vs. the identified nonlinear prediction. **d)** Power spectral density of the prediction error of Least Squares and PL-LSSVM.

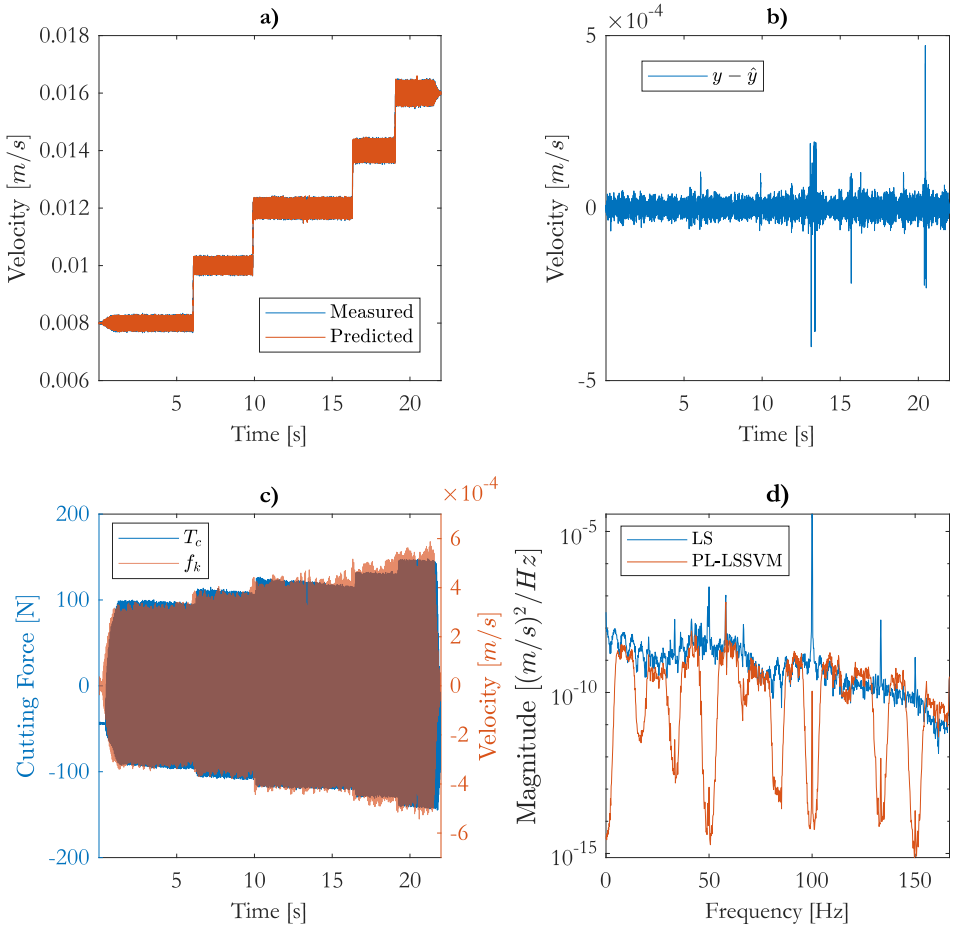

Figure 4.6: Results from the milling trial with increasing velocity trajectory from the milling case study using PL-LSSVM for system identification. **a)** Measured output of the system vs. the prediction by the PL-LSSVM model. **b)** Error between the measured output and the PL-LSSVM model prediction. **c)** Output from the endmill vs. the identified nonlinear prediction. **d)** Power spectral density of the prediction error of Least Squares and PL-LSSVM.

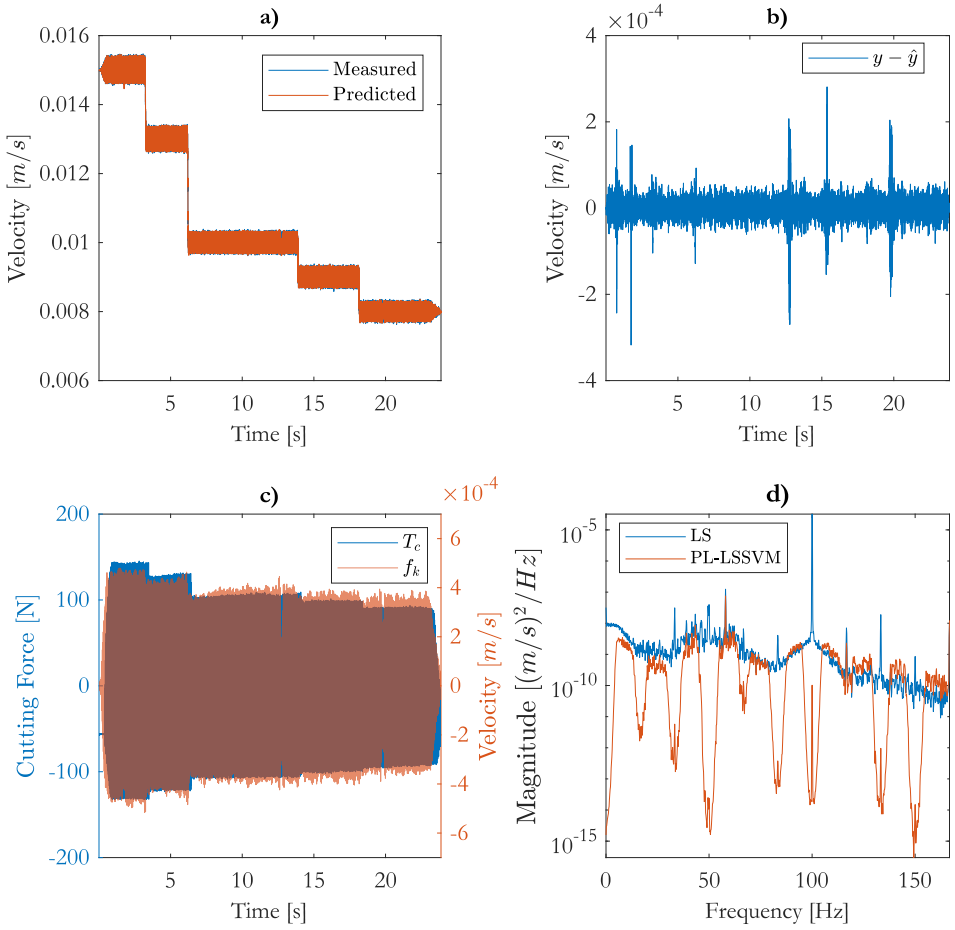

Figure 4.7: Results the milling trial with decreasing velocity trajectory from the milling case study using PL-LSSVM for system identification. **a)** Measured output of the system vs. the prediction by the PL-LSSVM model. **b)** Error between the measured output and the PL-LSSVM model prediction. **c)** Output from the endmill vs. the identified nonlinear prediction. **d)** Power spectral density of the prediction error of Least Squares and PL-LSSVM.

4.2 Kernel Recursive Least Squares - Tracker (KRLS-T)

4.2.1 Case Study #1: Shaker Test

The KRLS-T algorithm utilized the hand tuned hyperparameter values shown in in Table 4.5 along with the linear parameters identified from their respective trials by the PL-LSSVM method, shown previously in Table 4.2.

Hyperparameter	Value
σ_1	1.2
σ_2	0.99
γ	5.85
ψ	0.0182
λ	0.996

Table 4.5: Hyperparameters utilized in KRLS-T of Case Study #1

Figures 4.8, 4.9, and 4.10 show the results of the KRLS-T algorithm. The main detail of these figures is the similarity to the PL-LSSVM results in the figures 4.1, 4.2, and 4.3, in particular panels (a) and (c) which shows the prediction of the velocity and nonlinear disturbance. These results show that the KRLS-T predictions are close to the PL-LSSVM predictions. The initial prediction of KRLS-T is further off due to the lack of information available for prediction in the beginning.

The (b) panels illustrate the prediction error, $y - \hat{y}$, of the models created with parameters identified utilizing LS, PL-LSSVM and KRLS-T. The PSD of the PL-LSSVM and KRLS-T are more comparable than compared to the LS results for the two trials with disturbance. As discussed before the LS results show a large amount of prediction error at the disturbance frequency and its second harmonic in both Figures 4.9 and 4.10. Both kernel-based methods show a reduction in magnitude at the disturbance frequency though not equally. The PL-LSSVM shows a larger reduction than KRLS-T, especially in Figure 4.9. This shows that the KRLS-T velocity prediction does not contain the disturbance frequency component to the same extent as the PL-LSSVM method. To determine whether this is due to the kernel function we must investigate the nonlinear portion of the prediction.

To this aim, panel (d) displays the PSD of the measured force from the shaker and the disturbance in the velocity identified by PL-LSSVM and KRLS-T. The Figure 4.8 plot does not display any measured force as there was none measured from the shaker. For the trials with disturbance, both the PL-LSSVM and KRLS-T results display the same magnitude at the disturbance frequency and its harmonics. They diverge at other frequencies with the PL-LSSVM having a much

lower frequency content than KRLS-T. This demonstrates that both can identify disturbance frequency, but the KRLS-T also contains other frequencies. With regard to measured force, it is difficult to compare it directly against the predicted disturbance due to the difference in magnitude. Though of the two predicted disturbances the KRLS-T prediction is a better representation of the measured disturbance due to its inclusion of frequencies components besides the disturbance frequency.

Panel (e) describes this variance of the predicted cutting forces. Figure 4.8 illustrates the base variance for the trajectory since their is no disturbance during the trial. We see that the variance begins at 0 before rising and then leveling off. While the leveling takes time the range of the variance is small compared to the output. Figure 4.9 is the constant disturbance trial and its variance consists of a large spike when the disturbance begins, which then decreases as the kernel adapts before a second spike when the disturbance abruptly stops. Figure 4.10 is similar to the no disturbance test which shows that the kernel can track the sweeping disturbance without too much issue. The magnitude axis for the variance in that plot is larger than the no disturbance variance which does indicate that the kernel was more uncertain during this test. Overall, the variance for all trials is as expected and shows that the kernel function is more confident when predicting disturbances with less instantaneous changes.

Panel (f) details which basis elements were added and removed each iteration from the dictionary during the operation. The added element is always the current basis element while the element being removed varies. In most instances, the removed element is an older basis in the dictionary but there are instances where the recently added term is removed instead. To determine if this holds for actual cutting disturbances we will look at the second case study.

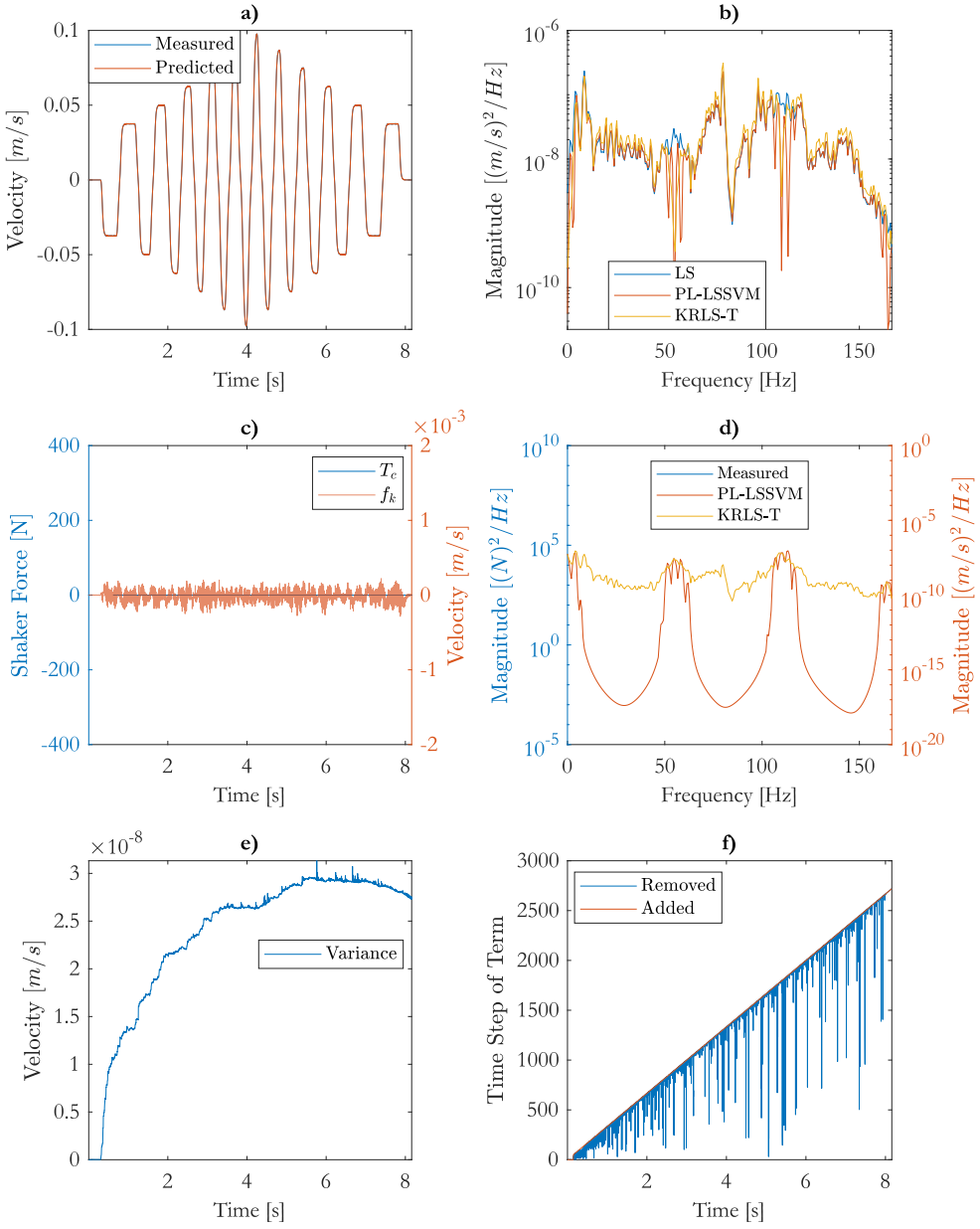

Figure 4.8: Results from the control trial of the shaker case study using KRLS-T for system identification and prediction. **a)** Predicted velocity, $\hat{y}$, versus the measured velocity of the feed drive, y. **b)** Power spectral density of the prediction error of LS, PL-LSSVM, and KRLS-T. **c)** Measured disturbance force from the shaker versus the predicted disturbance in the velocity. **d)** Power spectral density of the measured disturbance and disturbance predicted using PL-LSSVM and KRLS-T. **e)** The variance during the trial. **f)** The basis element is added and removed from the dictionary during the operation.

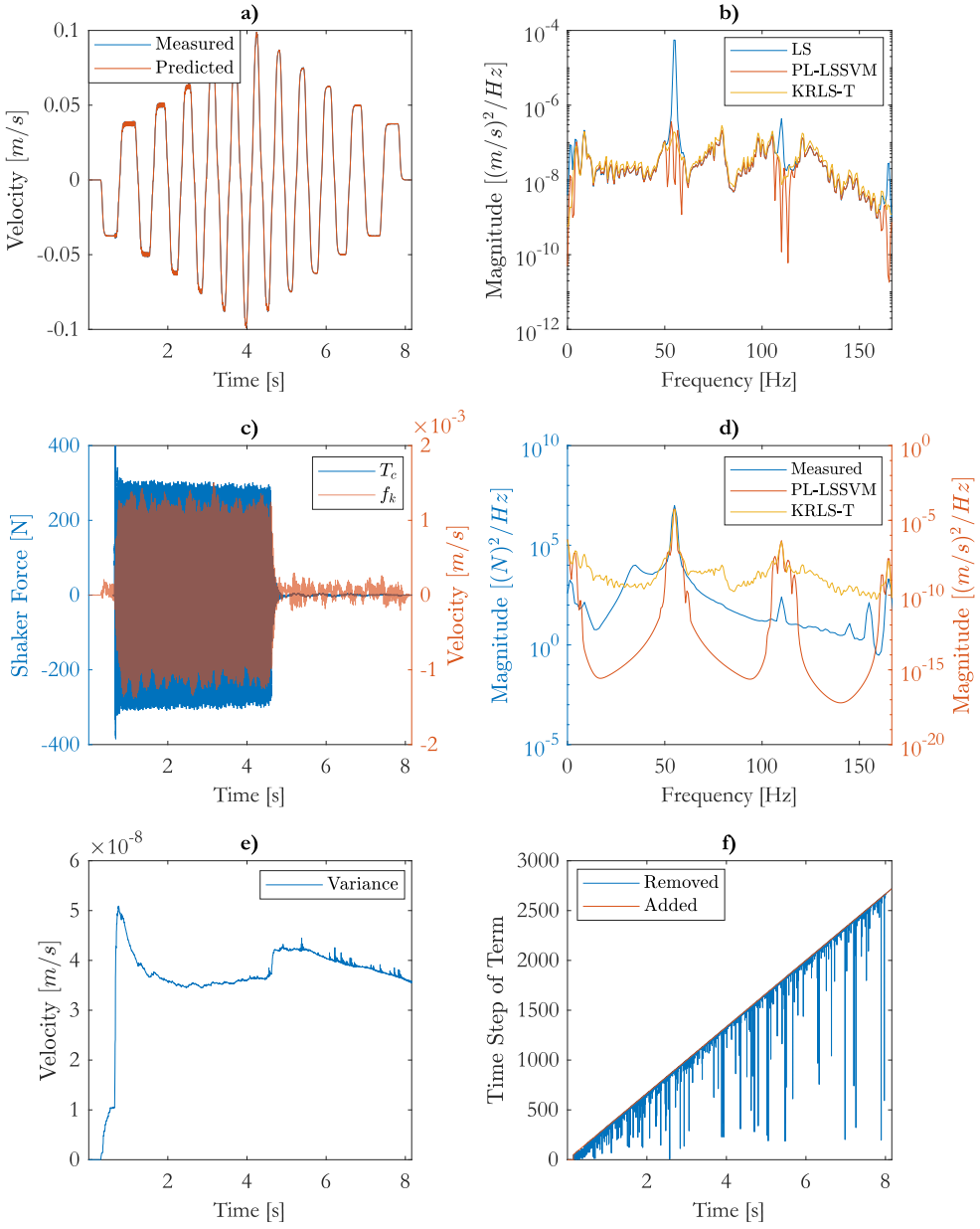

Figure 4.9: Results from the constant disturbance trial of the shaker case study using KRLS-T for system identification and prediction. **a)** Predicted velocity, $\hat{y}$, versus the measured velocity of the feed drive, y. **b)** Power spectral density of the prediction error of LS, PL-LSSVM, and KRLS-T. **c)** Measured disturbance force from the shaker versus the predicted disturbance in the velocity. **d)** Power spectral density of the measured disturbance and disturbance predicted using PL-LSSVM and KRLS-T. **e)** The variance during the trial. **f)** The basis element is added and removed from the dictionary during the operation.

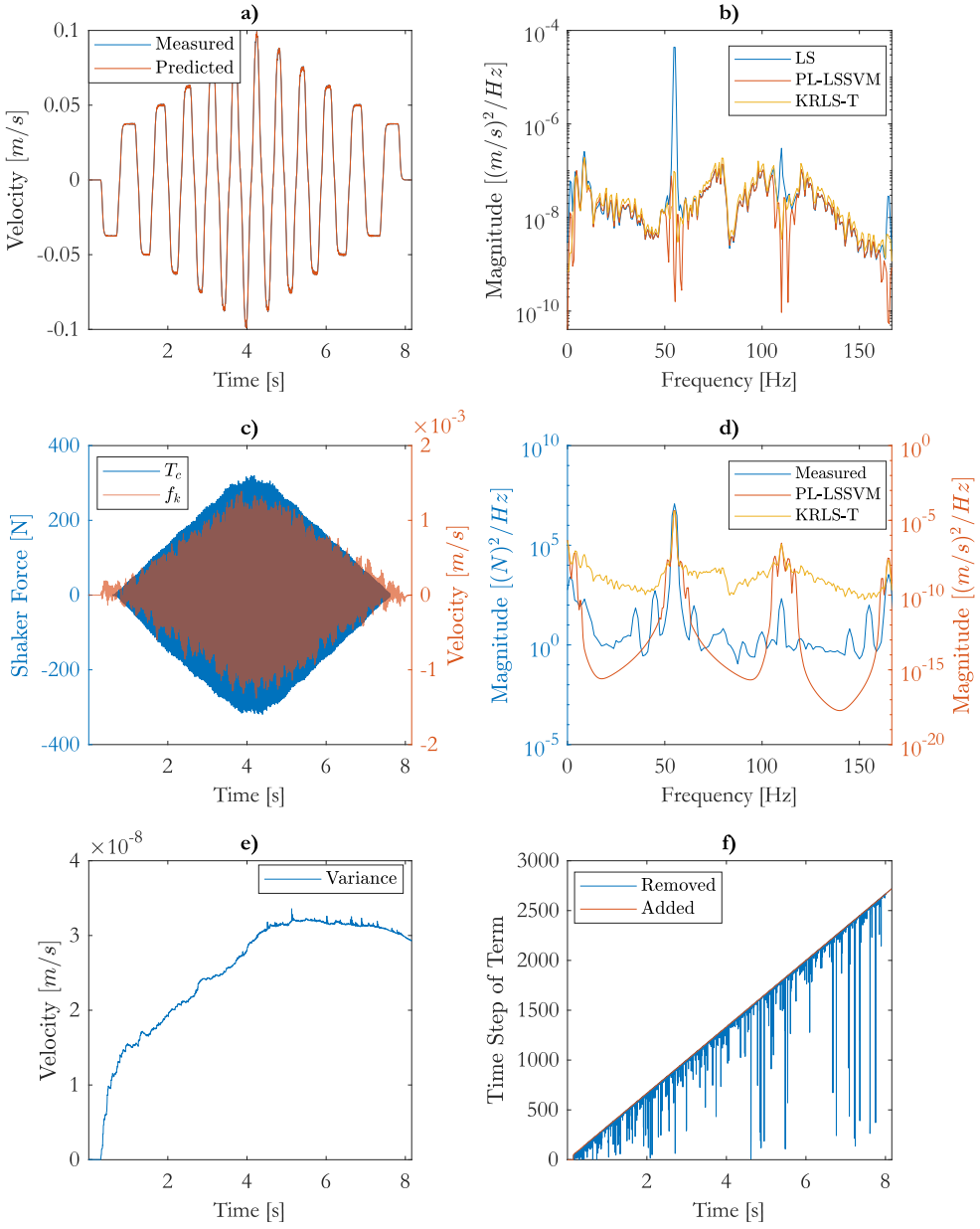

Figure 4.10: Results from the sweeping disturbance trial of the shaker case study using KRLS-T for system identification and prediction. **a)** Predicted velocity, $\hat{y}$, versus the measured velocity of the feed drive, y. **b)** Power spectral density of the prediction error of LS, PL-LSSVM, and KRLS-T. **c)** Measured disturbance force from the shaker versus the predicted disturbance in the velocity. **d)** Power spectral density of the measured disturbance and disturbance predicted using PL-LSSVM and KRLS-T. **e)** The variance during the trial. **f)** The basis element is added and removed from the dictionary during the operation.

4.2.2 Case Study #2: Milling Test

The KRLS-T algorithm utilized the hand tuned hyperparameter values shown in in Table 4.5 along with the linear parameters identified from their respective trials by the PL-LSSVM method, shown previously in Table 4.2. Unlike the first study where every trial had an identified set of parameters only the trial #1 and #2 had identified parameters. Therefore, the KRLS-T algorithm was provided the dynamic parameters identified from trial #1 for the KRLS-T application on trial #1 and the dynamic parameters from trial #2 were used for the KRLS-T of trials #2, #3, and #4. Figures 4.11, 4.12, 4.13, and 4.14 show the results of the KRLS-T algorithm for the four trials conducted in case study #2.

Hyperparameter	**Value**
σ_1	6.5
σ_2	0.92
γ	8
ψ	0.02
λ	0.997

Table 4.6: Hyperparameters utilized in KRLS-T of Case Study #2

Panel (a) and (c) illustrate the predicted versus measured velocity and force respectively. Panel (a) compares the measured velocity versus the velocity predicted by the KRLS-T algorithm. Comparing these results to the PL-LSSVM results shows that the results are similar, this makes sense as the linear parameters of the feed drive dynamics are the same for each trial with only the nonlinear disturbance from the milling forces being calculated differently. Panel (c) also shows similarities with the PL-LSSVM results including changes in the prediction due to the measurement errors discussed previously. From the predicted output the KRLS-T algorithm shows remarkably similar results to the PL-LSSVM.

Panel (b) shows the PSD of the prediction error for the LS, PL-LSSVM, and KRLS-T results. As seen previously in the results of the first case study, the KRLS-T can compensate for disturbance frequency though not to the same magnitude as the PL-LSSVM. Additionally, in this case study the PL-LSSVM has lower prediction error in other frequencies components. These frequencies are $1/3$ and $2/3$ of the disturbance frequency and its harmonics. This is due to interactions between the sampling rate of the controller, the chosen downsampling rate, and the period of the kernel.

Panel (d) illustrates the PSD of the measured force and predicted velocity disturbance by the PL-LSSVM and KRLS-T methods. In Figure 4.12, 4.13, and 4.14 we see that the shape of the

KRLS-T PSD is similar to the measured force. The main difference between the measured and KRLS-T disturbance are their overall magnitudes and the peak magnitude of the disturbance frequencies. The PL-LSSVM plot only contains significant magnitudes of the frequencies that it is shown to account for in Panel (b). This supports that the KRLS-T algorithm can predict the milling disturbance.

Panel (e) for Figure 4.11 illustrates the base variance for the case study. Figures 4.12, 4.13, and 4.14 all contain spikes in the variance which then decay over time. This is due to the measurements having periods where the output was not recorded and the measured value is the same as the last time step. This causes problems as the predictions from the KRLS-T algorithms can not account for this and cause momentary spikes in the variance which decays as the prior adapts to this information. Overall the magnitude of the variance is so small compared to the velocity output that it is inconsequential in both case studies.

Panel (f) details which basis elements were added and removed each iteration from the dictionary during the operation. The pruning operates every iteration once the dictionary has reached the size constraint, M, set before operation.

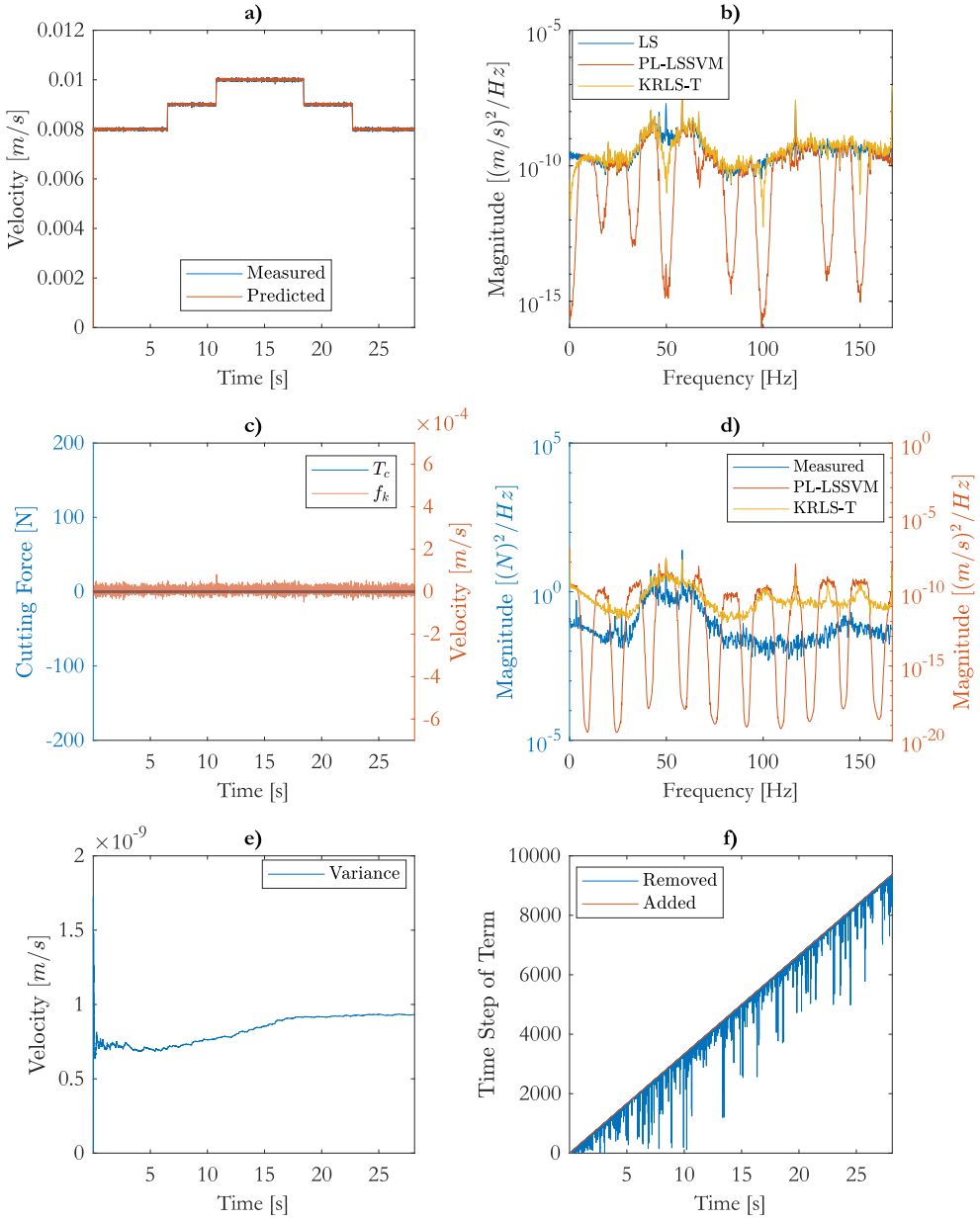

Figure 4.11: Results from the control trial of the milling case study using KRLS-T for system identification. **a)** Predicted velocity, $\hat{y}$, versus the measured velocity of the feed drive, y. **b)** Power spectral density of the prediction error of LS, PL-LSSVM, and KRLS-T. **c)** Measured disturbance force from the shaker versus the predicted disturbance in the velocity. **d)** Power spectral density of the measured disturbance and disturbance predicted using PL-LSSVM and KRLS-T. **e)** The variance during the trial. **f)** The basis element is added and removed from the dictionary during the operation.

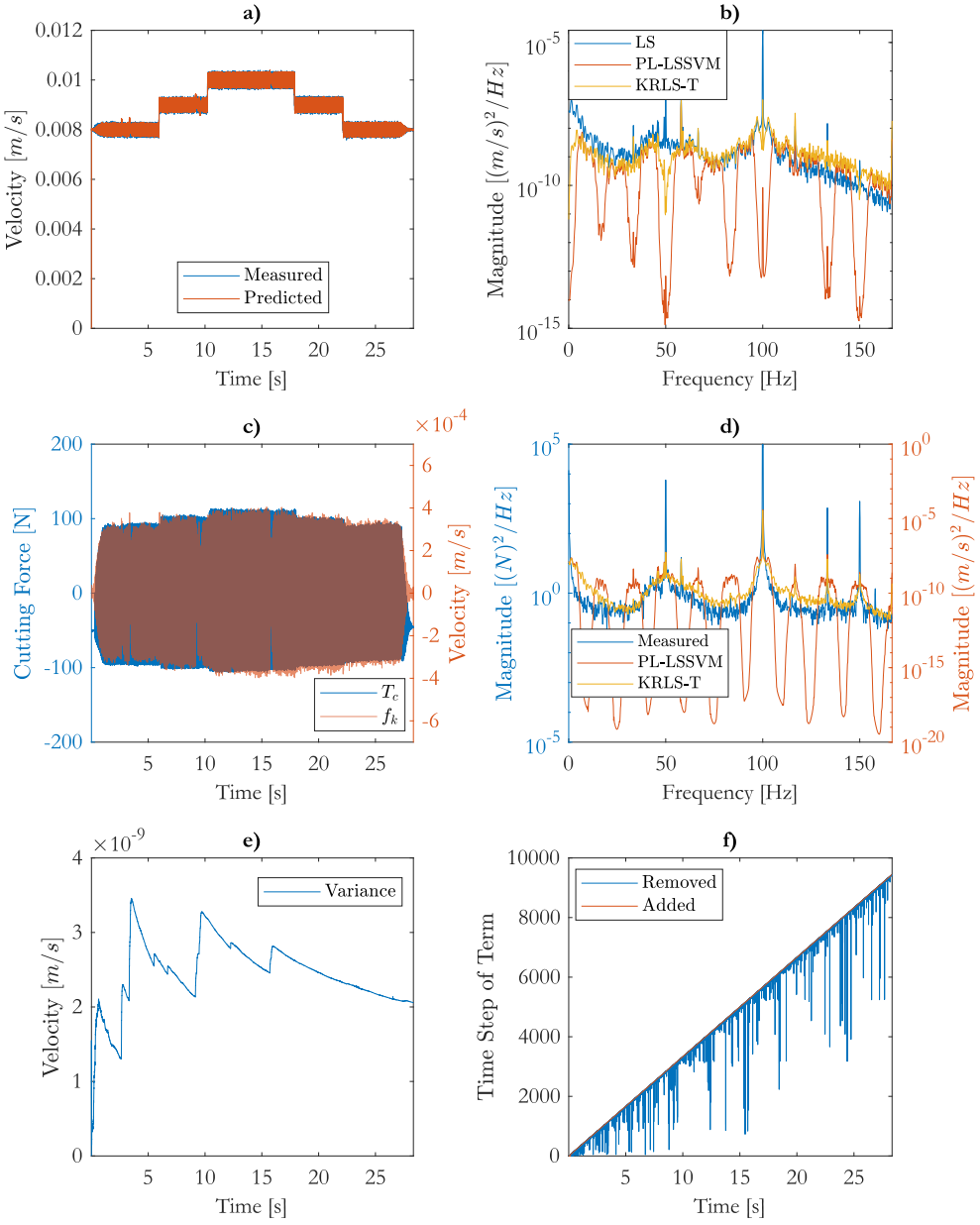

Figure 4.12: Results from the milling trial with increasing and decreasing velocity trajectory from the milling case study using KRLS-T for system identification. **a)** Predicted velocity, $\hat{y}$, versus the measured velocity of the feed drive, y. **b)** Power spectral density of the prediction error of LS, PL-LSSVM, and KRLS-T. **c)** Measured disturbance force from the shaker versus the predicted disturbance in the velocity. **d)** Power spectral density of the measured disturbance and disturbance predicted using PL-LSSVM and KRLS-T. **e)** The variance during the trial. **f)** The basis element is added and removed from the dictionary during the operation.

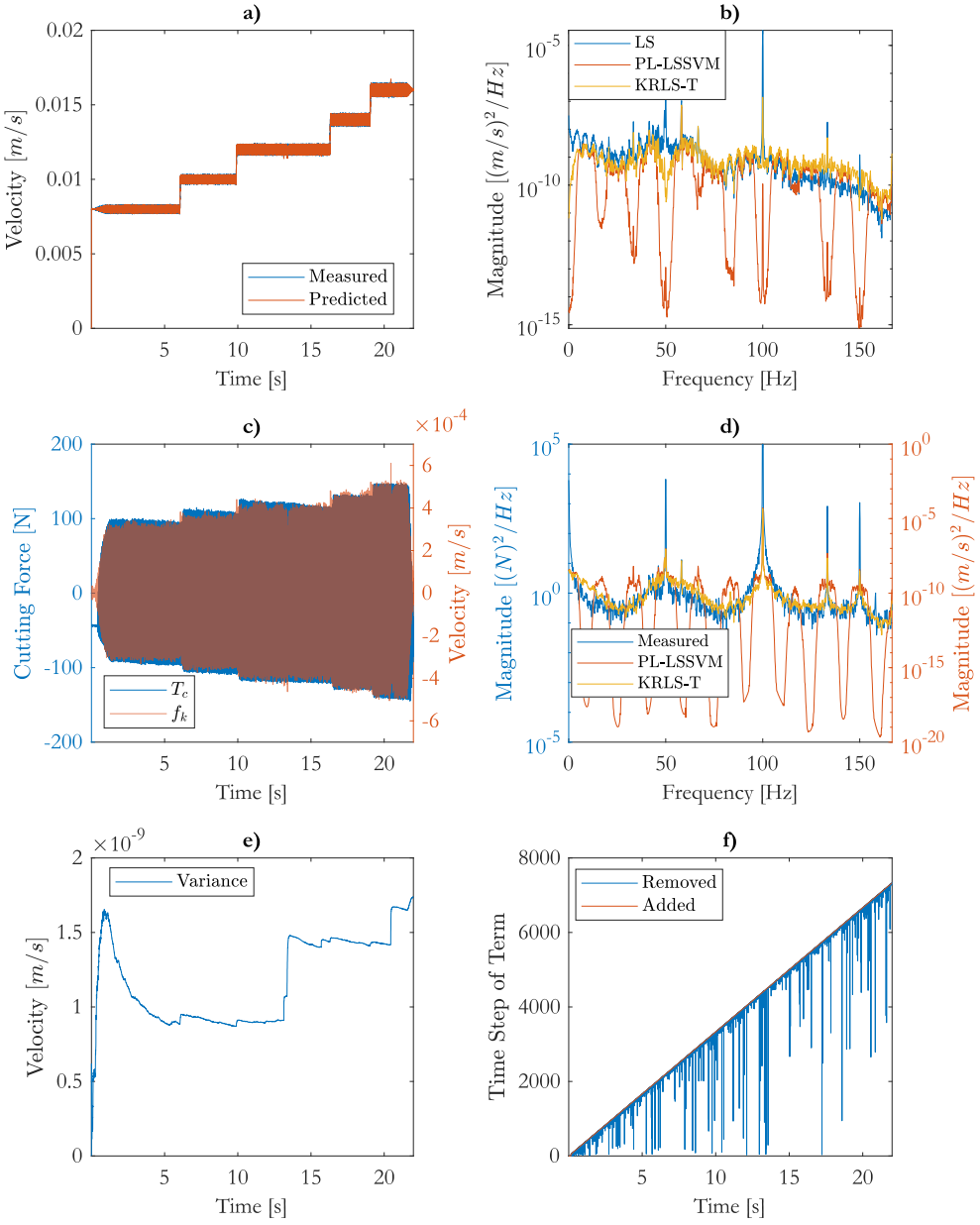

Figure 4.13: Results from the milling trial with increasing velocity trajectory from the milling case study using KRLS-T for system identification. **a)** Predicted velocity, $\hat{y}$, versus the measured velocity of the feed drive, y. **b)** Power spectral density of the prediction error of LS, PL-LSSVM, and KRLS-T. **c)** Measured disturbance force from the shaker versus the predicted disturbance in the velocity. **d)** Power spectral density of the measured disturbance and disturbance predicted using PL-LSSVM and KRLS-T. **e)** The variance during the trial. **f)** The basis element is added and removed from the dictionary during the operation.

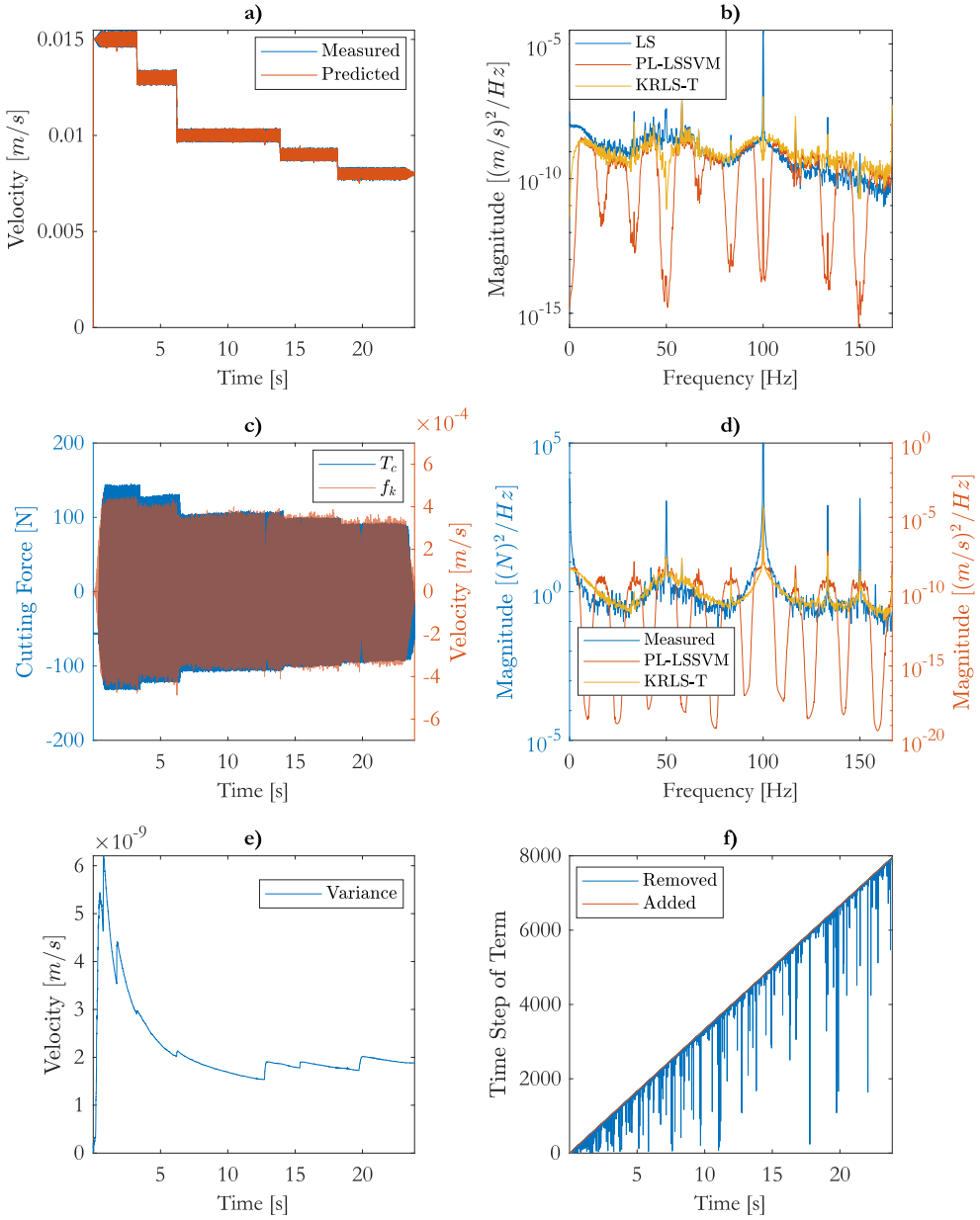

Figure 4.14: Results the milling trial with decreasing velocity trajectory from the milling case study using KRLS-T for system identification. **a)** Predicted velocity, $\hat{y}$, versus the measured velocity of the feed drive, y. **b)** Power spectral density of the prediction error of LS, PL-LSSVM, and KRLS-T. **c)** Measured disturbance force from the shaker versus the predicted disturbance in the velocity. **d)** Power spectral density of the measured disturbance and disturbance predicted using PL-LSSVM and KRLS-T. **e)** The variance during the trial. **f)** The basis element is added and removed from the dictionary during the operation.

4.3 Summary

This chapter covered the analysis of both case studies utilizing the PL-LSSVM for batch identification and the KRLS-T algorithm recursive disturbance prediction. These results demonstrate the ability of both methods to predict the cutting force bias from in-operation milling data, given that optimized hyperparameters are found beforehand. Comparisons between LS, PL-LSSVM, and KRLS-T are discussed to identify areas of similarity and difference between the results.

Chapter 5

Conclusions

Chapters 3 and 4 of this thesis investigated the ability of recursive kernel-based methods for online force prediction of cutting tools. The investigation that was conducted led to identifying the feed drive's dynamic parameters utilizing the batch kernel identification PL-LSSVM. The identified parameters of the feed drive's dynamics were used to remove the linear components of the dynamics from the recursive kernel-based algorithm. The KRLS-T algorithm was utilized to recursively predict the nonlinear disturbances from periodic forces acting on the feed drive. The main contribution of this work are summarized below.

5.1 Contribution

a) *Indirect measurement of feed drive disturbances from milling forces during online operation.*

The KRLS-T algorithm was adapted for online prediction of feed drive disturbances from milling forces during machining processes. The feed drive disturbances were predicted during online operation by a recursive kernel-based algorithm and the feed drive's dynamic model parameters identified offline previously. Comparing the results of the experimental evidences of the PL-LSSVM and KRLS-T algorithms verified the accuracy of the online operation.

5.2 Future Work

a) *Optimization of the hyperparameters for the kernel function, and PL-LSSVM and KRLS-T algorithms.*

The hyperparameters for the kernel function, and PL-LSSVM and KRLS-T algorithms

were selected after hand tuning the values for the measured results. These results confirm the application of these algorithms for indirect disturbance measurement but may not be the most optimal selection for this application. Therefore, one needs to devise a method for selecting a set of hyperparameters that optimizes the measurements.

b) *Recursive identification of the feed drive's dynamic model parameters.*

The online algorithm requires that parameters of the feed drive's dynamic model be identified before operation. Accurate identification of all the parameters requires proper excitation of the feed drive's dynamics within the data used for identification. Collecting this data within a single data set may not be applicable during typical online operation. Therefore, it would be practical to identify the parameters alongside the updating of the kernel dictionary during online operation.

Bibliography

[1] Y. Altintas and M. R. Khoshdarregi, "Contour error control of cnc machine tools with vibration avoidance," *CIRP annals*, vol. 61, no. 1, pp. 335–338, 2012.

[2] M. Tomizuka, "Zero phase error tracking algorithm for digital control," 1987.

[3] C. Okwudire, K. Ramani, and M. Duan, "A trajectory optimization method for improved tracking of motion commands using cnc machines that experience unwanted vibration," *CIRP Annals*, vol. 65, no. 1, pp. 373–376, 2016.

[4] M. Weck and G. Ye, "Sharp corner tracking using the ikf control strategy," *CIRP annals*, vol. 39, no. 1, pp. 437–441, 1990.

[5] K. Erkorkmaz and Y. Altintas, "High speed cnc system design. part ii: Modeling and identification of feed drives," *International Journal of Machine Tools and Manufacture*, vol. 41, no. 10, pp. 1487–1509, 2001.

[6] K. Erkorkmaz and Y. Altintas, "High speed cnc system design. part iii: High speed tracking and contouring control of feed drives," *International Journal of Machine Tools and Manufacture*, vol. 41, no. 11, pp. 1637–1658, 2001.

[7] G. Pritschow and W. Philipp, "Research on the efficiency of feedforward controllers in m direct drives," *CIRP annals*, vol. 41, no. 1, pp. 411–415, 1992.

[8] T. Bergs, D. Biermann, K. Erkorkmaz, and R. M'Saoubi, "Digital twins for cutting processes," *CIRP Annals*, vol. 72, no. 2, pp. 541–567, 2023.

[9] H.-C. Möhring, P. Wiederkehr, K. Erkorkmaz, and Y. Kakinuma, "Self-optimizing machining systems," *CIRP Annals*, vol. 69, no. 2, pp. 740–763, 2020.

[10] G. W. G. Tseng, C. Q. G. Chen, K. Erkorkmaz, and S. Engin, "Digital shadow identification from feed drive structures for virtual process planning," *CIRP Journal of Manufacturing Science and Technology*, vol. 24, pp. 55–65, 2019.

[11] C.-Y. Lee, S.-H. Hwang, E. Nam, and B.-K. Min, "Identification of mass and sliding friction parameters of machine tool feed drive using recursive least squares method," *The International Journal of Advanced Manufacturing Technology*, vol. 109, pp. 2831–2844, 2020.

[12] D. Aslan and Y. Altintas, "Prediction of cutting forces in five-axis milling using feed drive current measurements," *IEEE/ASME Transactions on Mechatronics*, vol. 23, no. 2, pp. 833–844, 2018.

[13] J. Zhu, T. Zhang, J. Wang, and X. Li, "Axial dynamic characteristic parameters identification of rolling joints in a ball screw feed drive system," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 230, no. 14, pp. 2449–2462, 2016.

[14] J. Zhang, J. Ding, N. Sugita, *et al.*, "A sample construction method in kinematics characteristics domain to identify the feed drive model," *Precision Engineering*, vol. 68, pp. 82–96, 2021.

[15] C. E. Okwudire and Y. Altintas, "Hybrid modeling of ball screw drives with coupled axial, torsional, and lateral dynamics," 2009.

[16] H. Ahmadian and H. Jalali, "Identification of bolted lap joints parameters in assembled structures," *Mechanical Systems and Signal Processing*, vol. 21, no. 2, pp. 1041–1050, 2007.

[17] Y. Altintas, A. Verl, C. Brecher, L. Uriarte, and G. Pritschow, "Machine tool feed drives," *CIRP annals*, vol. 60, no. 2, pp. 779–796, 2011.

[18] Q. Butler, Y. Ziada, D. Stephenson, and S. Andrew Gadsden, "Condition monitoring of machine tool feed drives: A review," *Journal of Manufacturing Science and Engineering*, vol. 144, no. 10, p. 100 802, 2022.

[19] C. C. De Wit, H. Olsson, K. J. Astrom, and P. Lischinsky, "A new model for control of systems with friction," *IEEE Transactions on automatic control*, vol. 40, no. 3, pp. 419–425, 1995.

[20] K. C. Cheok, H. Hu, and N. K. Loh, "Modeling and identification of a class of servomechanism systems with stick-slip friction," 1988.

[21] J.-H. Kim, H.-K. Chae, J.-Y. Jeon, and S.-W. Lee, "Identification and control of systems with friction using accelerated evolutionary programming," *IEEE Control systems magazine*, vol. 16, no. 4, pp. 38–47, 1996.

[22] C. C. De Wit and P. Lischinsky, "Adaptive friction compensation with partially known dynamic friction model," *International journal of adaptive control and signal processing*, vol. 11, no. 1, pp. 65–80, 1997.

[23] M. R. Popovic and A. A. Goldenberg, "Modeling of friction using spectral analysis," *IEEE Transactions on Robotics and Automation*, vol. 14, no. 1, pp. 114–122, 1998.

[24] B. Armstrong-Hélouvry, P. Dupont, and C. C. De Wit, "A survey of models, analysis tools and compensation methods for the control of machines with friction," *Automatica*, vol. 30, no. 7, pp. 1083–1138, 1994.

[25] B. Sencer and Y. Altintas, "Identification of 5-axis machine tools feed drive systems for contouring simulation," *International Journal of Automation Technology*, vol. 5, no. 3, pp. 377–386, 2011.

[26] M. Espinoza, J. A. Suykens, and B. De Moor, "Kernel based partially linear models and nonlinear identification," *IEEE Transactions on Automatic Control*, vol. 50, no. 10, pp. 1602–1606, 2005.

[27] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning* (Adaptive Computation and Machine Learning series), eng, 1st ed. Cambridge: MIT Press, 2005, ISBN: 9780262256834.

[28] G. Pillonetto, F. Dinuzzo, T. Chen, G. De Nicolao, and L. Ljung, "Kernel methods in system identification, machine learning and function estimation: A survey," *Automatica*, vol. 50, no. 3, pp. 657–682, 2014.

[29] Y. Engel, S. Mannor, and R. Meir, "The kernel recursive least-squares algorithm," *IEEE Transactions on signal processing*, vol. 52, no. 8, pp. 2275–2285, 2004.

[30] S. Van Vaerenbergh, M. Lázaro-Gredilla, and I. Santamaría, "Kernel recursive least-squares tracker for time-varying regression," *IEEE transactions on neural networks and learning systems*, vol. 23, no. 8, pp. 1313–1326, 2012.

[31] W. Liu, I. Park, Y. Wang, and J. C. Principe, "Extended kernel recursive least squares algorithm," *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3801–3814, 2009.

[32] Y. Liu, H. Wang, J. Yu, and P. Li, "Selective recursive kernel learning for online identification of nonlinear systems with narx form," *Journal of Process Control*, vol. 20, no. 2, pp. 181–194, 2010.

[33] S. Van Vaerenbergh, I. Santamaria, W. Liu, and J. C. Principe, "Fixed-budget kernel recursive least-squares," eng, in *2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2010, pp. 1882–1885, ISBN: 9781424442959.

[34] B. J. De Kruif and T. J. De Vries, "Pruning error minimization in least squares support vector machines," *IEEE Transactions on Neural Networks*, vol. 14, no. 3, pp. 696–702, 2003.

[35] L. Csató and M. Opper, "Sparse representation for gaussian process models," *Advances in neural information processing systems*, vol. 13, 2000.