IEEE 15th International Workshop on Advanced Motion Control
March 9-11, 2018, Tokyo, Japan

Kernel-Based Regression of Non-Causal Systems for Inverse Model Feedforward Estimation

Lennart Blanken
Department of Mechanical Engineering
Eindhoven University of Technology
Eindhoven, The Netherlands
l.l.g.blanken@tue.nl

Ids van den Meijdenberg
Department of Mechanical Engineering
Eindhoven University of Technology
Eindhoven, The Netherlands
j.w.a.v.d.meijdenberg@student.tue.nl

Tom Oomen
Department of Mechanical Engineering
Eindhoven University of Technology
Eindhoven, The Netherlands
t.a.e.oomen@tue.nl

***Abstract*—Inversion-based feedforward control enables high performance for industrial motion systems. To this end, accurate knowledge of the inverse system is required, and non-causal control actions must be enabled. The aim of this paper is to accurately identify non-causal inverse models in view of high feedforward control performance. The developed method employs kernel-based regularization to minimize the mean squared error of the estimate. The performance benefits of the presented approach are demonstrated on an industrial printing system, including non-causal feedforward control actions.**

***Index Terms*—Motion control; feedforward control; system identification; regularization; Gaussian process regression.**

I. INTRODUCTION

Feedforward can effectively compensate known disturbances before these affect the system, enabling perfect tracking. This is typically achieved by inverting the system through an inverse model-based feedforward, see, e.g., [1], [2], [3].

Inversion-based feedforward involves an identification step, either of the system followed by an inversion step, or through direct identification of the inverse system, see, e.g., [4]. The quality of the inverse model determines the control performance [5]: it should be capable to compensate for the excited dynamics between reference and tracking error. Especially for lightly damped mechatronic systems such as flatbed printers, see Fig. 3, constructing such accurate parametric models, and possibly inverting them, may be difficult and expensive, due to, e.g., complex dynamics [6], [7] and numerical issues [8]. This motivates to directly estimate the inverse system from measured input/output data, obtained in the same setting as the model is going to be used in [4], [6], [9].

Recent developments in identification for inversion-based feedforward are specifically directed towards closed-loop control performance with minimal variance. Based on the linear regression technique for low-order finite impulse response (FIR) model structures in [10], which was shown to yield biased estimates in [11], the use of instrumental variables (IV) is proposed in [11], [12] to obtain unbiased estimates with minimal variance. In view of compactly modeling rational (inverse) systems, these techniques have been extended towards the use of rational orthonormal basis functions with fixed poles [13], and rational output error (OE) model structures, see, e.g., [9], [14]. Interestingly, these developed models can be used to generate non-causal feedforward control actions, which are required to achieve high performance for non-minimum phase (NMP) systems. Yet, key points that remain are optimal selection of the model structure, e.g., low-order FIR or OE models, and the associated fixed model order in view of the well-known trade-off between bias and variance.

In the field of system identification, model order selection has received renewed interest, in particular by modeling systems as, possibly infinite dimensional, Gaussian processes [15], [16], [17]. Instead of restricting the model order through model order selection criteria such as AIC and cross-validation, see, e.g., [18], [19], in kernel-based regression the trade-off between bias and variance is handled by exploiting prior information. Seen from a Bayesian perspective, see, e.g., [17], [19], the impulse response of the system is hypothesized to be a realization of a zero-mean Gaussian stochastic process with some covariance function, called a kernel. The kernel can be used to encode priors on the impulse response, such as smoothness and exponential decay. Existing kernel structures include the stable spline kernel [16], diagonal/correlated kernel [17], and kernels based on rational orthonormal basis functions [20], [21], which in addition allow to express resonant behavior. This is particularly important for lightly damped systems, such as the flatbed printer in Fig. 3. The kernel essentially regularizes the estimates, and has been shown to drastically improve the bias/variance trade-off.

Although important progress has been made in identification of inverse systems for feedforward control, at present the advances in kernel-based identification have not yet been explored. This paper aims to develop a kernel-based identification procedure for non-causal systems, and apply this to inverse model estimation for feedforward control. Whereas kernel-based identification of *causal* systems in $\mathcal{RH}_2$ has received significant attention, see, e.g., [16], [17], [19], [20], here the focus is on the estimation of *non-causal* responses of inverse systems in $\mathcal{RL}_2$. Indeed, non-causal inverse models are required to achieve high tracking performance for non-minimum phase systems. The presented non-causal kernel is constructed based on non-causal rational orthonormal functions in $\mathcal{RL}_2$, see, e.g., [13]. This extends results on causal kernels based on rational orthonormal functions in $\mathcal{RH}_2$, see, e.g., [20], [21].

The contributions of this paper are threefold:

C1) A framework is developed for regularized identification of inverse systems for feedforward control;

C2) Development of a suitable non-causal kernel based on non-causal rational orthonormal functions, enabling high performance feedforward control;

C3) The benefits of the developed technique are demonstrated on an industrial flatbed printing system, including increased tracking performance and non-causal control.

Notation: To facilitate exposition, all systems are discrete-time, single-input single-output (SISO), and linear time-invariant. The complex indeterminate $z \in \mathbb{C}$ is omitted when this does not lead to any confusion. The following standard notation is used, see, e.g., [22]. Let $\mathbb{D}$ denote the open unit disc: $\{z, |z| < 1\}$, $\mathbb{T}$ the unit circle: $\{z, |z| = 1\}$, and $\mathbb{E}$ the complement of the closed unit disc: $\{z, |z| > 1\}$. $\mathcal{L}_2$ denotes the set of complex functions that are square integrable on $\mathbb{T}$, and the real rational subspace of $\mathcal{L}_2$ is denoted $\mathcal{RL}_2$. $\mathcal{H}_2$ denotes the set of complex functions that are square integrable on $\mathbb{T}$ and analytic for $|z| \geq 1$. $\mathbb{R}[z^{-1}]$ denotes the polynomial ring in indeterminate z^{-1} with coefficients in $\mathbb{R}$, and $\mathbb{R}[z, z^{-1}]$ denotes the Laurent polynomial ring in indeterminate z with coefficients in $\mathbb{R}$. Indeed, Laurent polynomials include both positive and negative exponents of the indeterminate. Signals are often tacitly assumed to be of length N. For a vector $x \in \mathbb{R}^N$, the squared two-norm is given by $\|x\|_2^2 = x^\top x$. W is positive definite if $x^\top W x > 0$, $\forall x \neq 0$.

II. PROBLEM DESCRIPTION

A. Feedforward Control Configuration

Consider the configuration in Fig. 1. The true unknown system is denoted $P(z)$. The control configuration consists of a given stabilizing feedback controller $C(z)$, and a feedforward controller $F(z)$ in serial arrangement. Let $r(t)$ denote a finite time reference with $t = 1, \ldots, N$, $y(t)$ the output, $\varepsilon(t)$ an i.i.d. zero-mean noise sequence uncorrelated with $r(t)$, and $\tilde{r}(t)$ the feedforward. The tracking error $e = r - y$ is given by

$$e(t) = (1 - T(q)F(q))\, r(t) - S(q)H(q)\varepsilon(t),$$

with $S(q) = (1 + P(q)C(q))^{-1}$ and $T(q) = S(q)P(q)C(q)$. Optimal tracking performance, i.e., $\overline{\mathbb{E}}\{e(t)\} = 0$, for all $r \neq 0$ is achieved if $F(z) = T^{-1}(z)$, where $\overline{\mathbb{E}}\{\}$ denotes mathematical expectation. Clearly, $T^{-1}(z)$ is unknown.

B. Problem Formulation

The aim of identification for feedforward is to estimate the inverse system $T^{-1}(z)$, based on data $\{r(t), y(t)\}_{t=1}^{N}$. Essentially, the pursued approach corresponds to identification of parametric model $F(q, \theta)$, given *input* data $y(t)$ and *output* data $r(t)$ from the open-loop data generating system

$$r(t) = T^{-1}(q)y(t) - v_u(t),$$

where $v_u(t) = T^{-1}(q)S(q)H(q)\varepsilon(t)$. This interpretation is depicted in Fig. 2. To simplify the analysis, it is assumed that $v_u(t)$ is i.i.d. zero-mean white noise with variance σ^2.

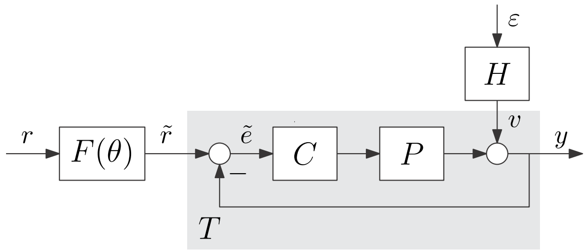
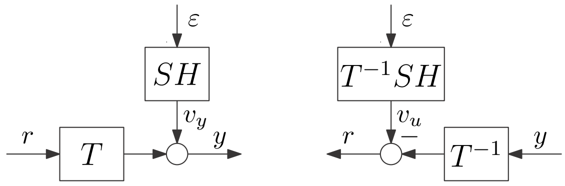


Fig. 1. Closed-loop control configuration with serial feedforward.



(a) Forward dynamics. (b) Backward dynamics.

Fig. 2. Forward and backward identification setting.

Important aspects in identification for feedforward are dealing with stability of the inverse system, and selection of the model structure and order. Indeed, if $T(z)$ has non-minimum phase dynamics, i.e., zeros in $\mathbb{E}$, $T^{-1}(z)$ has unstable poles. For feedforward however, this does not pose a problem, since unstable poles can be accounted for through non-causal filtering operations. This is illustrated in the next section using an industrial motion system. The developed kernel-based regression techniques for feedforward control, see Sections IV and V, are then applied to this system in Section VI.

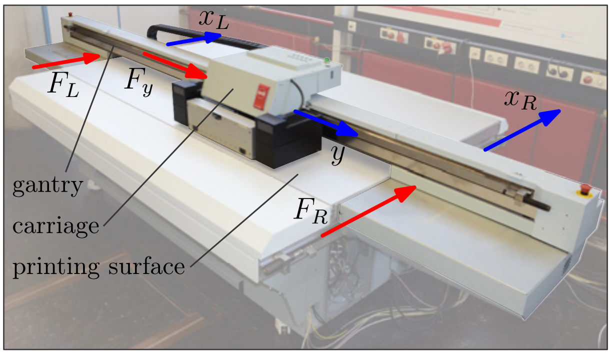
Remark 1. *Alternative to the serial configuration, $F(q, \theta)$ can be placed parallel to $C(q)$, see, e.g., [11], where additional attention is paid to reveal closed-loop identification issues.*

III. INVERSE SYSTEM IDENTIFICATION FOR FEEDFORWARD: DEALING WITH NMP DYNAMICS

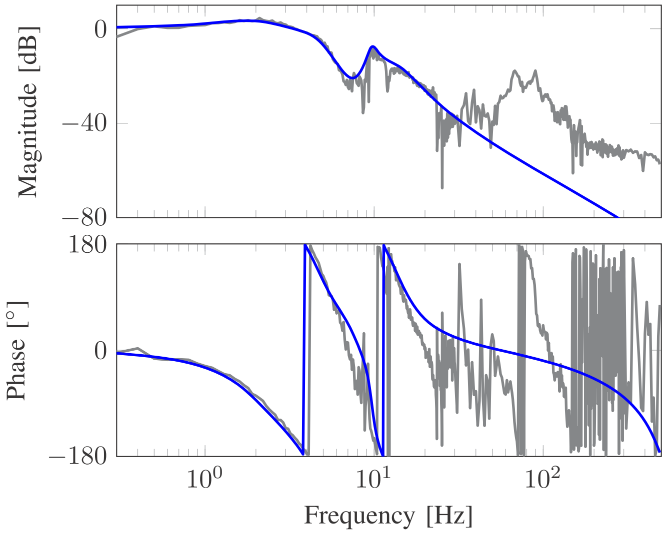
A. Case Study: Industrial Flatbed Printer

Consider the Océ Arizona 550 GT flatbed printer, shown in Fig. 3. In contrast to standard printers, the medium is fixed on the printing surface. The carriage, which contains the printheads, moves in the horizontal plane. The system is equipped with three current-driven brushless electric motors to provide the required forces F_L, F_R, F_y. The measurement system consists of three optical encoders, collocated with the actuators, yielding position measurements x_L, x_R, y. A stabilizing multivariable feedback controller is implemented. The system is operated with sampling time 10^{-3} s.

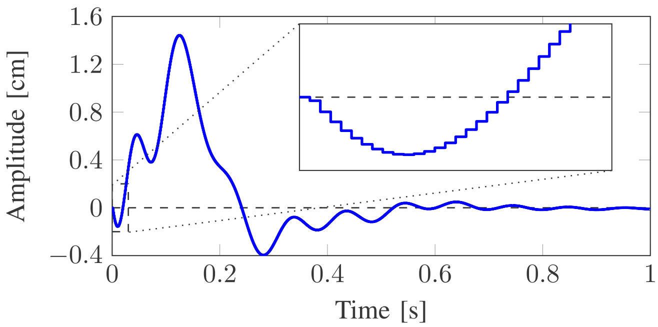
The input and output considered for control are the current u_L [A] to the actuator on the left side of the gantry, and the position x_R [m] on the right side of the gantry. By closing all feedback loops, a SISO equivalent system $P : u_L \to x_R$ is obtained. Since all performed displacements are small, the system can be assumed linear. An identified frequency response function of the closed-loop $T = \frac{PC}{1+PC}$ is shown in Fig. 3b. Due to the non-collocated input/output pairing, the



(a) Photograph of Océ Arizona 550 GT flatbed printer.



(b) Bode diagram of FRF measurement (——) and model (——) of T.



(c) Impulse response of model of T.

Fig. 3. Océ Arizona 550 GT flatbed printing system, with actuator forces in red and sensor locations in blue. For simulations, the system P from input F_L to non-collocated output x_R is considered. The NMP dynamics of T can be observed in its impulse response, which initially moves in negative direction.

system exhibits non-minimum phase dynamics. This can be observed from its impulse response, which initially moves in negative direction. A 10^{th}-order model of T is derived for simulations in Section VI, containing three NMP zeros.

B. Stability of Inverse System

For systems T with NMP dynamics, the optimal feedforward $F = T^{-1}$ has poles in $\mathbb{E}$. Often, T^{-1} is interpreted as a causal and unbounded operator on $\mathcal{H}_2$. In contrast, here T^{-1} is interpreted as a bounded and non-causal operator on $\mathcal{L}_2$, see also [23, Section 1.5] in the context of feedback control. Using the bilateral Z-transform, the inverse system can be described by the rational representation with infinite impulse response

$$T^{-1}(z) = \frac{B_0(z)}{A_0(z)} = \sum_{k=-\infty}^{\infty} \theta_k^{\text{o}} z^{-k}, \qquad (1)$$

with $A_0(z), B_0(z) \in \mathbb{R}[z^{-1}]$. The non-causal impulse response $\{\theta_k^{\text{o}}\}$ arises from solving the underlying difference equation also for negative time, e.g., with boundary conditions of zero at $\pm\infty$. The non-causal and bounded response in (1) allows to uniquely compute bounded feedforward signals $\tilde{r}$ that yield optimal tracking, i.e., $\overline{\mathbb{E}}\{e(t)\} = 0$. Next, a non-causal model structure for F is introduced in view of (1), and the estimation is formulated as a linear regression problem.

C. Non-Causal FIR Estimation for Feedforward Control

The following non-causal model $F(q, \theta)$ is considered:

$$F(q, \theta) = \sum_{k=-n_{ac}}^{n_c} \theta_k q^{-k} = \Psi(q)\theta \qquad (2)$$

where n_{ac} is the number of anti-causal terms, n_c the number of strictly causal terms, and $n_\theta = n_{ac} + n_c + 1$. The basis functions $\Psi(q) = [q^{n_{ac}}, q^{n_{nc-1}}, \ldots, q^0, \ldots, q^{-n_c}]$ are Laurent polynomials, i.e., $q^{-k} \in \mathbb{R}[q, q^{-1}]$, as often used in feedforward control, see, e.g., [11]. This is a generalization of the widely used FIR models in system identification [18], [19], [20], yet allows for using preview in feedforward control.

Since (2) is linear in the parameters, the estimation problem can be written as a linear regression model

$$r_N = \Phi_N \theta + v_N,$$

where $r_N \in \mathbb{R}^N$ is the data, $\Phi_N = \Psi(q)y(t) \in \mathbb{R}^{N \times n_\theta}$ is the regression matrix, and v_N is the white Gaussian disturbance distributed as $\mathcal{N}(0, \sigma^2 I_N)$. Using a standard linear regression method, see, e.g., [18], [24], the estimates for the backward identification setting can then be obtained using least squares:

$$\hat{\theta}_{LS} = \arg\min_{\theta} \|r_N - \Phi_N \theta\|_2^2 = \left(\Phi_N^\top \Phi_N\right)^{-1} \Phi_N^\top r_N. \quad (3)$$

To accurately model (1), high complexity models in terms of n_θ may be required. This inherently leads to large variance of $\hat{\theta}_{LS}$, which may result in poor tracking performance. To reduce estimation errors, e.g., in mean squared sense, existing approaches often focus on trimming the model order, at the cost of introducing additional bias. Instead in the next section, a kernel-based approach is presented for feedforward.

IV. KERNEL-BASED REGULARIZED FIR ESTIMATION FOR FEEDFORWARD CONTROL

A kernel-based approach is developed to estimate optimal non-causal feedforward controllers, constituting contribution C1. The approach builds on results on kernel-based regression in the field of system identification, see, e.g., [17], [19], yet in contrast focuses on estimating non-causal models.

The parameters are estimated by regularized least squares:

$$\hat{\theta}_{ReLS} = \arg\min_{\theta} \|r_N - \Phi_N\theta\|_2^2 + \sigma^2\theta^\top D^{-1}(\alpha)\theta \quad (4)$$

$$= \left(D(\alpha)\Phi_N^\top\Phi_N + \sigma^2 I_N\right)^{-1} D(\alpha)\Phi_N^\top r_N, \quad (5)$$

where $D(\alpha) \succeq 0$ is the regularization matrix, or kernel matrix, and α are hyperparameters. In case $D(\alpha)$ is singular, (4) should be interpreted as in [25, Remark 2.1]. Since the kernel D regularizes parameters θ corresponding to both anti-causal and causal terms in (2), it is referred to as a non-causal kernel.

Remark 2. *The method* (4) *admits a Bayesian interpretation, see [17], [19]. Indeed, $\{\theta^o\}$ is hypothesized to be a realization of a zero-mean Gaussian process with prior covariance D. Given data $\{r(t), y(t)\}_{t=1}^N$, the conditional distribution is $\theta|r_N \sim \mathcal{N}(\hat{\theta}_{ReLS}, D^{post})$, with posterior mean $\hat{\theta}_{ReLS}$.*

In view of tracking performance, the used kernel D in (4) plays a key role. It is shown in [17, Theorem 1] that the optimal kernel with respect to the mean square error (MSE) matrix of the estimate $\hat{\theta}_{ReLS}$, defined as

$$MSE(\hat{\theta}_{ReLS}, D) = \overline{\mathbb{E}}\big(\hat{\theta}_{ReLS} - \theta^o\big)\big(\hat{\theta}_{ReLS} - \theta^o\big)^\top,$$

is given by $D^{\text{opt}} = \theta^o\theta^{o\top}$, in the sense that for any $D \succeq 0$, it holds $MSE(\hat{\theta}_{ReLS}, D) \succeq MSE(\hat{\theta}_{ReLS}, D^{\text{opt}})$. Clearly, the optimal kernel D^{opt} is unknown in practice. Yet, it provides a guideline for design of D. Priors on the system can be used to choose a structure $D(\alpha)$ in terms of hyperparameters α. Relevant dynamic properties include resonances, i.e., complex pole pairs, and exponential decay of the (non-causal) impulse response. Especially for mechatronic systems, such information may well be available from FRF measurements, which are often accurate and inexpensive to obtain [26]. In the next section, a non-causal kernel is developed that is suited for feedforward control of lightly damped mechatronic systems.

V. A NON-CAUSAL KERNEL BASED ON ORTHONORMAL BASIS FUNCTIONS IN $\mathcal{L}_2$

In this section, a non-causal parametric kernel structure is developed that is suited for estimation of non-causal inverse systems for feedforward control, constituting contribution C2. The kernel exploits and builds on non-causal rational orthonormal basis functions (ROBFs) in $\mathcal{RL}_2$, see, e.g., [13], to allow for lightly damped non-causal responses. This is in sharp contrast with causal kernels that are developed in the system identification community, see, e.g., [20], [21], which are based on ROBFs in $\mathcal{RH}_2$, see, e.g., [22], [27].

First a frequency-domain construction approach is presented for non-causal orthonormal basis functions. Then, these functions are used to construct time-domain non-causal kernels.

A. Non-Causal Orthonormal Basis Functions in $\mathcal{L}_2$

A construction approach is presented for the non-causal rational orthonormal basis functions. The orthonormality is with respect to the standard inner product on $\mathcal{L}_2$:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \psi_k(e^{i\omega})\psi_l(e^{-i\omega})d\omega = \delta_{k,l} = \begin{cases} 1 & \text{if } k = l, \\ 0 & \text{if } k \neq l. \end{cases}$$

The considered orthonormal basis functions are defined by the sequences $\xi_s = \{\xi_{s,k}\}_{k=1,2,\ldots}$ and $\xi_u = \{\xi_{u,k}\}_{k=0,-1,\ldots}$, with $\xi_{s,k}, \xi_{u,k} \in \mathbb{D}$, $\forall k$, which define the poles of the rational basis functions. The basis functions are given by

$$\psi_k(z) = \begin{cases} \frac{\sqrt{1-|\xi_{s,k}|^2}}{z-\xi_{s,k}}\phi_k(z,\xi_s), & \text{if } k > 0, \\ \frac{\sqrt{1-|\xi_{u,k}|^2}}{1-\overline{\xi_{u,k}}z}\phi_k'(z,\xi_u), & \text{if } k \leq 0, \end{cases} \quad (6)$$

where the all-pass transfer functions ϕ_k, ϕ_k' are defined by

$$\phi_k(z,\xi) = \begin{cases} 1 & \text{if } k = 1, \\ \prod_{m=1}^{k-1} \frac{1-\overline{\xi_m}z}{z-\xi_m} & \text{if } k > 1. \end{cases}$$

$$\phi_k'(z,\xi) = \begin{cases} 1 & \text{if } k = 0, \\ \prod_{m=1}^{k-1} \frac{z-\xi_m}{1-\overline{\xi_m}z} & \text{if } k < 0. \end{cases}$$

The sequence $\{\psi_k\}_{k>0} \subseteq \mathcal{H}_{2-}$ forms the well-known Takenaka-Malmquist functions, and consists of strictly causal functions. The set $\{\psi_k\}_{k\leq 0} \subseteq \mathcal{H}_{2-}^{\perp}$ contains anti-causal functions and direct feedthrough terms, e.g., select $\xi_{u,0} = 0$. The space $\mathcal{H}_{2-}$ denotes all functions in $\mathcal{H}_2$ that are zero at infinity, such as strictly proper systems. $\mathcal{H}_{2-}^{\perp}$ denotes the orthogonal complement of $\mathcal{H}_{2-}$ in $\mathcal{L}_2$. To ensure real-valued impulse responses, unitary transformations of (6) are required [27], and the poles ξ_s and ξ_u should occur in complex pairs.

B. Non-Causal Kernels Based on Orthonormal Functions

To construct the kernel, the functions (6) are transformed to the time domain through the (inverse) Fourier transform:

$$\psi_k(e^{i\omega}) = \mathcal{F}\{\varphi_k(t)\}, \quad \varphi_k(t) = \mathcal{F}^{-1}\{\psi_k(e^{i\omega})\},$$

where $\mathcal{F}$ denotes the Fourier transform, defined as $X(\omega) = \sum_{t=-\infty}^{\infty} x(k)e^{-j\omega t}$. Then, the space spanned by basis functions $\{\varphi_k(t)\}_{k=-n_u}^{n_s}$, denoted $\mathcal{H}_D$ in the sequel, is a reproducing kernel Hilbert space (RKHS) with reproducing kernel

$$D(t,t') = \textstyle\sum_{k=-n_u}^{n_s} \varphi_k(t)\varphi_k(t'), \quad -n_{ac} \leq t, t' \leq n_c. \quad (7)$$

In other words, the kernel D induces an hypothesis space $\mathcal{H}_D$ in which the impulse response to be estimated is expected to lie. By including the regularization term $\theta^\top D^{-1}\theta$, the estimate $\hat{\theta}_{ReLS}$ in (4) is regularized towards the hypothesis space $\mathcal{H}_D$.

Remark 3. *The developed non-causal kernel* (7) *recovers the causal kernel based on orthonormal basis functions as developed in, e.g., [20], [21], as a special case by choosing all poles of the basis functions* (6) *inside $\mathbb{D}$, i.e., $\xi_u = \emptyset$.*

C. Hyperparameter Selection

The poles of basis functions (6) can be treated as hyperparameters α of the kernel (7). In view of D^{opt}, these are ideally chosen close to the poles of T^{-1}, i.e., the zeros of T.

The hyperparameters for kernel-based regression are often estimated by the empirical Bayes approach [20], [21] in terms of marginal likelihood maximization, see, e.g., [15]. For mechatronic systems, knowledge on the zeros of the system may well be available a priori, for instance based on inexpensive non-parametric FRF measurements [26]. Hence, the hyperparameters may also be set manually by the user.

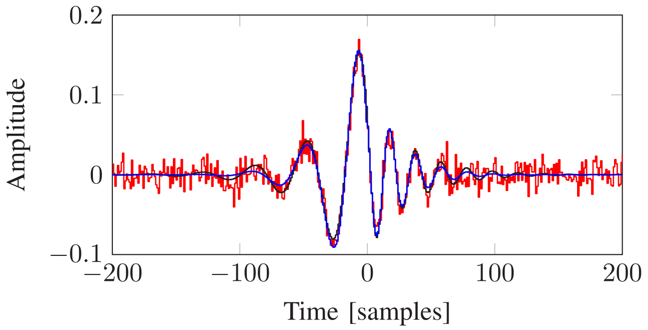


Fig. 4. Results case I: estimated impulse responses. The estimates $\hat{\theta}_{LS}$ without regularization (—) exhibit high variance. Using regularization (—), the MSE $\|\hat{\theta}_{ReLS} - \theta^o\|_2^2$ is significantly reduced: variance is traded for some bias with respect to true impulse response (—).

VI. EXAMPLES

In this section, forming contribution C3, the developed approach for kernel-based regression of inverse systems for feedforward control is validated using two case studies:

- Case I: the kernel-based identification approach using non-causal kernel (7) is validated on an example system;
- Case II: the benefits for feedforward control of the kernel-based identification approach are demonstrated on the industrial flatbed printer, see Subsection III-A.

A. Case I: Identification of Non-Causal Fourth Order System

The validity of the regularized identification approach is illustrated on an artificial example system, including non-causal regularization. Consider the continuous-time system

$$T^{-1}(s) = \prod_{i=1}^{2} \frac{\omega_i^2}{s^2 + 2\zeta_i \omega_i s + \omega_i^2},$$

with natural frequencies $\omega_1 = 1$ [Hz], $\omega_2 = -0.5$ [Hz] and damping ratios $\zeta_1 = 0.1$, $\zeta_2 = 0.2$. The discrete-time system $T^{-1}(z) \in \mathcal{RL}_2$ is obtained using zero-order-hold with sampling time $T_s = 0.02$ [s], and has one complex pole pair at $0.922 \pm 0.298i$ in $\mathbb{D}$, and one pole pair at 1.020 ± 0.158 in $\mathbb{E}$. Its non-causal impulse response is shown in Fig. 4.

Two 401[st] order non-causal FIR models, see (2), are estimated with $n_{ac} = n_c = 200$: one *without* regularization according to (3), and one *with* regularization, see (5), using non-causal kernel (7). The kernel is defined by poles ξ_s and ξ_u, which are chosen equal to the true poles of $T^{-1}(z)$, except have twice as much damping and are repeated twice:

$$\{\xi_s\} = \{0.895 \pm 0.285i, 0.895 \pm 0.285i\},$$
$$\{\xi_u\} = \{1.054 \pm 0.153i, 1.054 \pm 0.153i\}.$$

The estimation data is generated as $r(t) = T^{-1}(q)y(t) - v_u(t)$ with $N = 1000$, where $y(t)$ and $v_u(t)$ are uncorrelated i.i.d. zero-mean white noise sequences with variances $\sigma_y^2 = 1$ and $\sigma^2 = 0.3$, respectively. The results are shown in Figures 4 and 5, and the following observations are made:

- The non-regularized estimates $\hat{\theta}_{LS}$ suffer from high variance, which can also be observed from $F(\hat{\theta}_{LS})$.

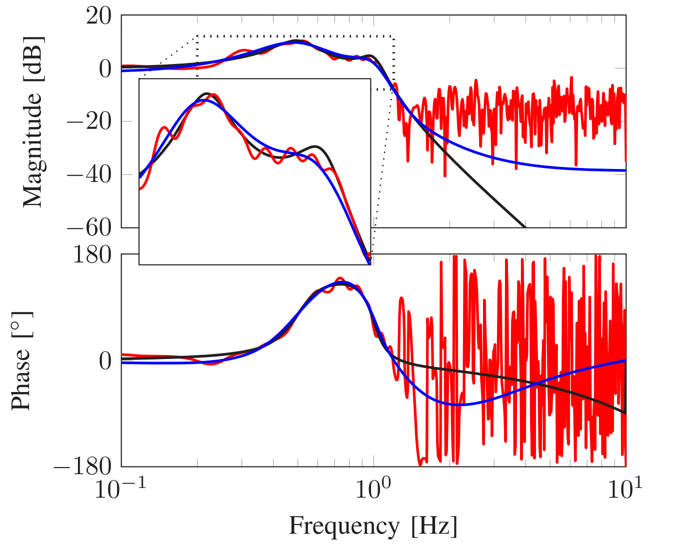


Fig. 5. Results case I: Bode diagram of true inverse system T^{-1} (—), $F(\hat{\theta}_{LS})$ (—) and $F(\hat{\theta}_{ReLS})$ (—). The decreased variance of $\hat{\theta}_{ReLS}$ results in a smoother frequency response, also at high frequencies where the signal-to-noise ratio is extremely low (approximately −23 dB at 2 Hz).

- Using regularization with non-causal kernel (7), the trade-off between bias and variance is drastically improved. This results in a smoother response $F(e^{i\omega}, \hat{\theta}_{ReLS})$.

B. Case II: Feedforward Control of Industrial Flatbed Printer

Next, the benefits for feedforward control of the developed approach are demonstrated on the Océ Arizona flatbed printer, see Subsection III-A and Fig. 3, which exhibits NMP dynamics. Two tasks $r(t)$ are performed, with $N = 3000$, see Fig. 6.

- First, the estimation data is generated as $y(t) = T(q)r(t)$, where the measured $r(t)$ is contaminated with i.i.d. zero mean white noise with variance $\sigma^2 = 10^{-3}$, see Fig. 2b.
- Second, the estimated feedforward controllers are implemented: $y(t) = T(q)F(q)r(t)$, see Fig. 1, without noise to compare only based on the identification approach.

Two 1001[st] order non-causal FIR models are estimated with $n_{ac} = n_c = 500$: one *without* regularization (3) and one *with* regularization (5) using kernel (7). The poles ξ_s and ξ_u are chosen equal to the poles of $T^{-1}(z)$. The results are depicted in Figures 6 and 7. The following observations are made:

- The estimates without regularization have high variance, resulting in erratic feedforward signals that are undesirable in practice, and relatively poor tracking performance.
- The developed approach using non-causal regularization leads to improved performance: the feedforward signal closely matches the optimal feedforward, and $\|e\|_2$ is reduced by 50% compared to the non-regularized case.
- To effectively compensate for the non-minimum phase dynamics of T, non-causal feedforward is applied.

Summarizing, the results confirm the effectiveness of the developed kernel-based regression approach for inversion-based feedforward control, including improved non-causal control.

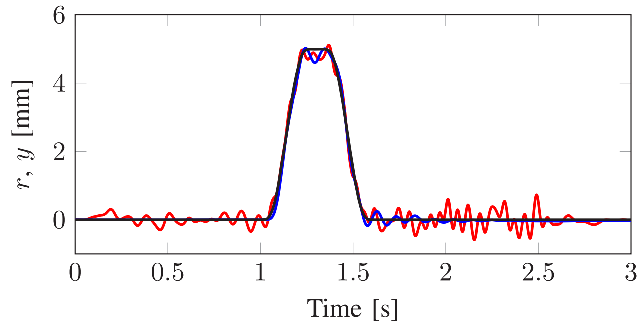


Fig. 6. Results case II (Arizona printer): tracking performance with $F(\hat{\theta})$ implemented. Without regularization (—), significant errors remain due to high variance of the estimates. The developed regularized approach leads to considerably improved tracking (—) of reference signal r (—).

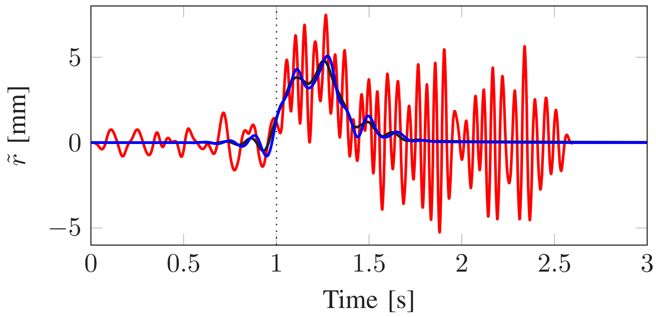


Fig. 7. Results case II (Arizona printer): feedforward signals $\tilde{r} = F(\hat{\theta})r$. Using regularization (—) leads to smooth feedforward signal that closely matches the optimal feedforward $T^{-1}r$ (—). Without regularization (—), the high variance of the estimates leads to erratic control actions. To compensate for the NMP dynamics of T, non-causal control is used: feedforward is applied before the motion task starts at $t = 1$ s.

VII. CONCLUSIONS

A kernel-based regression approach is developed to identify inverse systems for feedforward control. At present, feedforward control-oriented identification approaches often involve estimation of low-order models to avoid variance errors, at the cost of potential performance. The main contribution of this paper is the development of a non-causal kernel-based regularization approach based on non-causal orthonormal functions, which enables increased performance for non-minimum phase systems. The benefits of the developed approach are demonstrated on an industral printer, including non-causality.

ACKNOWLEDGMENT

This work is supported by Océ Technologies, and is part of the research programme VIDI with project number 15698, which is (partly) financed by the Netherlands Organisation for Scientific Research (NWO).

REFERENCES

[1] G. M. Clayton, S. Tien, K. Leang, Q. Zou, and S. Devasia, "A review of feedforward control approaches in nanopositioning for high-speed SPM," *Journal of Dynamic Systems, Measurements, and Control*, vol. 131, no. 6, p. 061101, 2009.

[2] K. Saiki, A. Hara, K. Sakata, and H. Fujimoto, "A study on high-speed and high-precision tracking control of large-scale stage using perfect tracking control method based on multirate feedforward control," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 4, pp. 1393–1400, 2010.

[3] J. van Zundert and T. Oomen, "On inversion-based approaches for feedforward and ILC," *IFAC Mechatronics*, 2017.

[4] Y. Jung and M. Enqvist, "Estimating models of inverse systems," in *Proc. of the 52nd IEEE Conf. on Decision and Control*, Florence, Italy, 2013, pp. 7143–7148.

[5] S. Devasia, "Should model-based inverse inputs be used as feedforward under plant uncertainty?" *IEEE Transactions on Automatic Control*, vol. 47, no. 11, pp. 1865–1871, 2002.

[6] M. Gevers, "Identification for control: From the early achievements to the revival of experiment design," *European Journal of Control*, vol. 11, no. 4, pp. 335 – 352, 2005.

[7] H. Hjalmarsson, "System identification of complex and structured systems," *European Journal of Control*, vol. 15, no. 3, pp. 275 – 310, 2009.

[8] R. van Herpen, O. Bosgra, and T. Oomen, "Bi-orthonormal polynomial basis function framework with applications in system identification," *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3285–3300, 2016.

[9] L. Blanken, F. Boeren, D. Bruijnen, and T. Oomen, "Batch-to-batch rational feedforward control: from iterative learning to identification approaches, with application to a wafer stage," *IEEE/ASME Transactions on Mechatronics*, vol. 22, no. 2, pp. 826–837, 2017.

[10] S. van der Meulen, R. Tousain, and O. Bosgra, "Fixed structure feedforward controller design exploiting iterative trials: Application to a wafer stage and a desktop printer," *Journal of Dynamic Systems, Measurements, and Control*, vol. 130, no. 5, pp. 0 510 061–05 100 616, 2008.

[11] F. Boeren, T. Oomen, and M. Steinbuch, "Iterative motion feedforward tuning: A data-driven approach based on instrumental variable identification," *Control Engineering Practice*, vol. 37, pp. 11 – 19, 2015.

[12] F. Boeren, D. Bruijnen, and T. Oomen, "Enhancing feedforward controller tuning via instrumental variables: with application to nanopositioning," *International Journal of Control*, vol. 90, no. 4, pp. 746–764, 2017.

[13] L. Blanken, G. Isil, S. Koekebakker, and T. Oomen, "Flexible ILC: Towards a convex approach for non-causal rational basis functions," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 12 107–12 112, 2017, IFAC 2017 World Congress, Toulouse, France.

[14] F. Boeren, L. Blanken, D. Bruijnen, and T. Oomen, "Optimal estimation of rational feedforward control via instrumental variables: With application to a wafer stage," *Asian Journal of Control*, 2017, to appear.

[15] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. The MIT Press, 2006.

[16] G. Pillonetto and G. D. Nicolao, "A new kernel-based approach for linear system identification," *Automatica*, vol. 46, no. 1, pp. 81 – 93, 2010.

[17] T. Chen, H. Ohlsson, and L. Ljung, "On the estimation of transfer functions, regularizations and Gaussian processes-Revisited," *Automatica*, vol. 48, no. 8, pp. 1525 – 1535, 2012.

[18] L. Ljung, *System identification - Theory for the User*. Upper Saddle River, NY, USA: Prentice-Hall, 1999.

[19] 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.

[20] T. Chen and L. Ljung, "Regularized system identification using orthonormal basis functions," in *Proc. of the 2015 European Control Conf. (ECC)*, Linz, Austria, 2015, pp. 1291–1296.

[21] M. Darwish, G. Pillonetto, and R. Tóth, "The quest for the right kernel in bayesian impulse response identification: the use of OBFs," *Automatica*, vol. 87, pp. 318 – 329, 2018.

[22] P. Heuberger, P. Van den Hof, and B. Wahlberg, *Modelling and Identification with rational orthogonal basis functions*. Springer-Verlag, London, UK, 2005.

[23] G. Vinnicombe, *Uncertainty and Feedback: $\mathcal{H}_\infty$ loop-shaping and the ν-gap metric*. London, U.K.: Imperial College Press, 2001.

[24] T. Söderström and P. Stoica, *System identification*. Hemel Hempstead, U.K.: Prentice Hall, 1989.

[25] T. Chen, M. S. Andersen, L. Ljung, A. Chiuso, and G. Pillonetto, "System identification via sparse multiple kernel-based regularization using sequential convex optimization techniques," *IEEE Transactions on Automatic Control*, vol. 59, no. 11, pp. 2933–2945, 2014.

[26] R. Pintelon and J. Schoukens, *System identification: a frequency domain approach*. John Wiley & Sons, 2012.

[27] B. Ninness and F. Gustafsson, "A unifying construction of orthonormal bases for system identification," *IEEE Transactions on Automatic Control*, vol. 42, no. 4, pp. 515–521, 1997.