A suppress-excite approach for online trajectory generation of uncertain motion systems

Mechanical Systems and Signal Processing 186 (2023) 109769

Contents lists available at ScienceDirect

Mechanical Systems and Signal Processing

journal homepage: www.elsevier.com/locate/ymssp

A suppress-excite approach for online trajectory generation of uncertain motion systems

Yazan M. Al-Rawashdeh [a], Mohammad Al Janaideh [a,*], Marcel F. Heertjes [b]

[a] Department of Mechanical Engineering, Memorial University, St. John's, Newfoundland A1B 3X5, Canada
[b] ASML and Department of Mechanical Engineering at Eindhoven University of Technology, 5612 AZ Eindhoven, Netherlands

ARTICLE INFO

Communicated by J.E. Mottershead

Keywords:
Trajectories
Flexible modes
Suppress
Excite
Input shaping
Uncertain
Real-time
Point-to-point
Closed-form

ABSTRACT

We present a framework that allows designing point-to-point closed-form trajectories and identifying flexible modes, jointly. The impulse response of a k-cascaded second-order notch filter is altered by composing it with a polynomial function of even degree in the time domain. The filter impulse response parameters are designed based on the available prior knowledge about the lightly-damped uncertain modes of the driven stage. The resulting chirps signals suppress the response of these oscillatory modes within selected suppression bands in the frequency domain. To check for unknown flexible modes within any desired frequency intervals of interest, the complement of the suppression bands is excited using excitation chirps signals. Using transmissibility, then the frequencies associated with the unknown flexible modes are identified, and the k-cascaded notch filter is updated accordingly. This update process may involve optimizing the parameters of the original filter, or the new filter when the order k changes. Also, real-time trajectory extrapolation is guaranteed under the proposed framework by the simple calculations required. The effectiveness of the proposed framework is illustrated through a numeric simulation example where the integration of both kinematic and dynamic constraints gives rise to the notion of kinodynamical trajectories.

1. Introduction

Various complex systems that are used in modern machining, material handling and manufacturing centers contain precision positioning systems. For example, in wafer scanners [1] which are influenced by the increased demand of higher throughput and yield, precision motion must be achieved under high acceleration profiles, which are 180 m/s^2 for the reticle stage and 45 m/s^2 for the wafer stage, during scanning motion [2,3]. The resulting inertial forces induce undesired vibrations in the lithography tool [3], that typically cause reticle-slip phenomena [4]. Consequently, these vibrations prevent the lithography tool from reaching the desired critical dimension (CD) and the specifications on overlay. Moreover, the induced vibrations can affect the optics system and cause leveling errors which lead the physical location of the silicon wafer to be out of the lens focal plane [2]. A similar challenge caused by structural vibrations is encountered by computer numerical controlled (CNC) machines used in machining [5], which even gets worse when the process-induced vibrations [6] are due to material removal frictional forces.

Triggering lightly-damped modes by the inertial forces can induce structural vibrations. These vibrations can be reduced by the motion control itself [5], active vibration suppression or isolation techniques [1], input shaping [7–9] and filtering techniques [10]. Avoid triggering these modes in the first place seems reasonable [5], and therefore it is recommended to look into the making of the reference trajectories before applying any other methods that either compensate, suppress or mask the effect of these structural

* Corresponding author.
E-mail addresses: yalrawashdeh@mun.ca (Y.M. Al-Rawashdeh), maljanaideh@mun.ca (M. Al Janaideh), marcel.heertjes@asml.com (M.F. Heertjes).

https://doi.org/10.1016/j.ymssp.2022.109769
Received 3 June 2022; Received in revised form 23 August 2022; Accepted 4 September 2022
Available online 20 October 2022

vibrations on the closed-loop servo (error) signals. Prior to this, however, the lightly damped modes triggered by the inertial forces should be modeled or identified using a suitable method [11]. Also, parametric uncertainties and bounded variation, or/and unmodeled dynamics should be considered. In this study, we consider mainly the parametric uncertainties responsible for rendering the lightly-damped modes uncertain, e.g., modal mass, stiffness, and damping. These uncertainties are addressed in the form of frequency bands to be *avoided* while generating the reference signal.

Based on the application, repetitive [12] and non-repetitive point-to-point (P2P), curvilinear, and free-form motion trajectories can be designed using, for example, Bezier, B-spline and non-uniform rational basis spline (NURBS) [13], and linear interpolation [14]. By definition, the NURBS is a geometric representation of the desired trajectory and is not system-aware, i.e. not aware of the kinematical and dynamical limitations of the uncertain flexible motion system. Therefore, an interpolation process is needed to effectively realize NURBS trajectories [15,16]. Usually, curvilinear and free-form trajectories are approximated using sequences of tiny P2P linear motions, which can then be made system-aware [14].

For example, the desired system-aware reference trajectories can be designed using input shapers [17], trapezoidal [18], trigonometric functions, and S-curve input smoothing [19], continuous-time polynomials [5,20,21], polynomial basis functions [8], templates [22], finite-impulse response (FIR) filters [10,12,14], or filters based on infinite-impulse responses (IIR) [23]. According to [24], using input shapers results in faster and more effective vibration reduction in flexible systems compared to input smoothing, [18,19]. In [17], zero-vibration (ZV) and zero-vibration-derivative (ZVD) were achieved using mainly input shapers with time delays of various distribution. A ZVD shaper results from two ZV shapers in cascaded. The idea behind a ZV shaper is to locate pairs of the shaper infinitely many zeros at the oscillatory modes, which can be done in time and frequency domains. At each point in a P2P trajectory, a system-aware continuous-time polynomial [5], requires minimizing its discrete Fourier spectrum, and the determination of its constant coefficients while considering the kinematical constraints. Therefore, closed-form solutions are difficult to obtain, resulting in limited trajectory extrapolation capabilities. To overcome the extrapolation limitation of generic polynomials, rational polynomial basis functions can be used, however, the resulting solutions lack useful analytical forms. In [8], an iterative feedforward control scheme – that comprises joint input shaping and feedforward parameterization – using polynomial basis functions was successful in extrapolating repetitive P2P trajectories using iterative analytical expressions. To ensure the system safety, the obtained parameterization have to be checked against given physical bounds. Interestingly, linear and nonlinear systems input commands can be generated using templates [22], where the desired input is convolved with certain impulse sequence — usually obtained from the driven system impulse response. Using such templates, a unity magnitude zero vibration (UMZV) can be obtained, which is faster than ZV input shapers and can be used with bang–bang control systems [25,26], and extra-insensitive (EI) input shaper that is more robust to frequency modeling errors than the ZVD shaper.

In essence, input shapers mostly depend on convolution which can be defined as "a mathematical operation on two functions of an input function and a convoluted function, producing an output function that is typically viewed as a modified version of input function" [27]. In the field of trajectory generation for precision control, convolution extensively finds its way through FIR-based trajectory generation techniques [10,12,14,27], while it is not recommended being used with IIR filters – like notch filters – because these alter both the amplitude and the phase of a signal around the notch frequency in a way that distorts the reference trajectory [5,9] such that the target position is not achieved, or a delay is introduced. Using successive convolution operations to accommodate physical system kinematical limits result in stretching the total travel time of the designed trajectories [10], which is not suitable for time-critical applications, e.g. wafer scanners, because it will reduce the system throughput [18]. Combining, for example, the ideas from [10,27] will result in system-aware trajectories whose robustness against flexible modes variations or parametric uncertainties is yet to be established. Moreover, under such a combination, increased number of flexible modes or even the increase in their frequencies will be problematic. Therefore, it seems reasonable to look for other alternatives that may provide reduced – preferably zero – vibration like ZV, robustness similar to ZVD and EI shapers, management of the needed travel time, ability to accommodate large number of uncertain flexible modes with minimum travel time, trajectory extrapolation capabilities, and closed-form solutions, all while being fully system-aware. Hence, we have this investigation.

Both convolution and function composition are two means to create new functions or signals with desired features. Function composition is a non-commutative mathematical process that – *in this study* – takes two functions $N_T(t) : \mathbb{R} \to \mathbb{R}$ and $g(t) : \mathbb{R} \to \mathbb{R}$ to produce another function $N_T(g(t)) \equiv (N_T \circ g)(t) : \mathbb{R} \to \mathbb{R}$. According to the herein proposed framework, when the information about the frequency ranges of lightly-damped modes – found in a flexible and uncertain linear time-invariant (LTI) system – is given, then it is mapped into a suitable frequency-domain template as a transfer function, say $N_T(s)$. Usually, the desired motion trajectories are given in the time domain where any kinematic constraints can be defined easily. To realize these kinematic constraints, the motion trajectories are usually designed by handling the acceleration, or other higher-order derivatives in the time domain to produce a suitable time-domain template, say $g(t)$. To simultaneously match the kinematic and dynamic requirements, i.e. to create a kinodynamical trajectory [28], we propose composing both the frequency-domain and time-domain templates in time domain to produce $N_T(g(t))$ that inherits the properties of both templates. This of course requires finding $N_T(t)$, i.e., the impulse response, by taking the inverse Laplace transform of $N_T(s)$ using appropriate techniques [29,30]. By carefully designing both templates, analytical expressions of the resulting point-to-point (P2P) trajectories can be provided and used to extrapolate the relative motion between *target-points* in real-time through simple calculations. At the target point, the velocity, acceleration, and preferably other higher-order derivatives must be zero. This is mainly the essence of this study. For *way-points*, the kinematical constraints may be nonzero, in general.

In FIR-based trajectory generation under convolution, FIR filters are applied in series, resulting in stretching the overall travel time by the sum of the total delay in the chain [10]. Therefore, the more flexible modes to suppress, the more stretching in the travel time to be expected; since each FIR filter corresponds to one frequency to be suppressed. Compared to that, the proposed

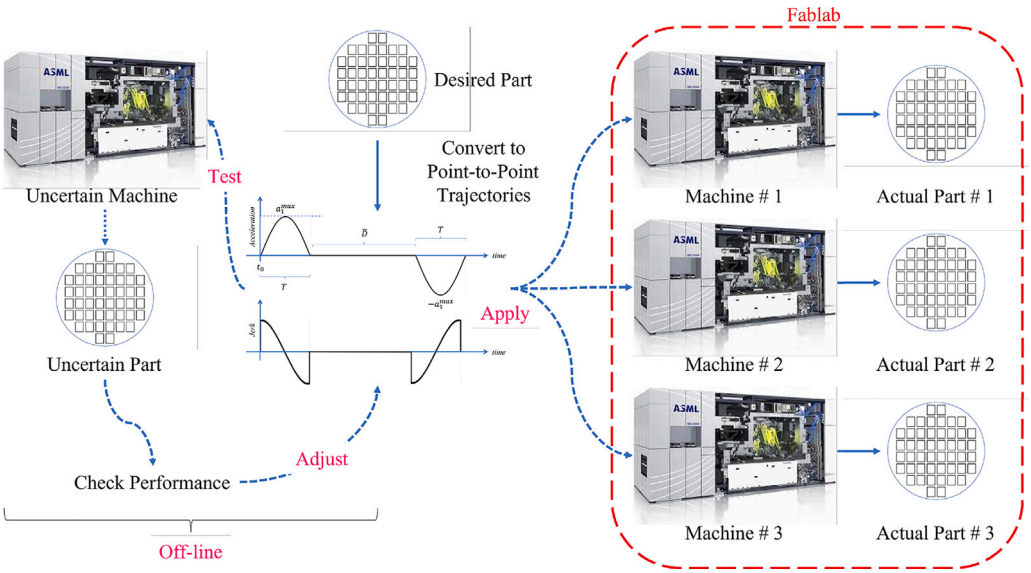

Fig. 1. An example from semiconductor manufacturing describing the problem at hand were three – *supposedly identical* – wafer scanners in a fablab produce a desired part. Due to machine variation, the produced parts may vary significantly as well. So, these machines are modeled and lumped into one uncertain machine model, and the produced uncertain part is checked against a set of performance metrics which may demand an off-line update of the used point-to-point trajectories. This update of the designed trajectories will reduce the deviation of the actual parts produced from the desired part.

function composition approach allows introducing as many as desired of flexible modes related frequencies using cascaded IIR notch filters while controlling the overall travel time through manipulating the involved parameters. Also, dealing with uncertain flexible modes, and detecting unknown flexible modes while in motion are made possible under the proposed framework. To keep this paper focused, a detailed comparison between the two approaches will be covered in a separate work.

To the best of the authors' knowledge, function composition is not *yet* used in trajectory generation for precision motion, in contrast to other fields. For example, in the field of volume rendering, the selection of the appropriate sampling rate of a function composition between a function representing a transfer function, i.e. impulse response, assigning opacities to data given by another scalar function is covered in [31]. In this study, function composition is used to generate reference trajectories that are meant for precision motion applications. Compared to convolution-based methods, the herein proposed method is not only capable of suppressing any number of known lightly-damped modes without stretching the total travel time [10], but also capable of identifying unknown lightly-damped modes while in motion.

To assist the reader through this paper, the following narrative is suggested: a machine builder who builds- supposedly exact-replicas of positioning stages wonders if it is possible to have a generally applicable and tunable reference trajectory profile that can steer all stages- despite any existing differences among them- with a specified performance without altering the *standard* control loops used. Therefore, to address such a problem, we use an uncertain linear time-invariant dynamical model to capture the common dynamics of all stages where the design or manufacturing variations are handled as *known* uncertainties. We also consider any stage-dependent dynamics as *unknown* flexible modes that may deteriorate the stage performance. Having these flexible modes known, then the reference trajectory is tuned accordingly to retain the desired performance of any stage while keeping the control loops fixed. An example from semiconductor manufacturing is depicted in Fig. 1 where the wafer scanner time-domain modeling is detailed in [32].

Based on the previous problem description, the proposed framework can be used under two themes, i.e., dealing with open/closed-loop plant variation under guaranteed performance, and optimal (time or throughput) trajectory generation for uncertain plants under guaranteed performance. In this investigation, we focus on the first theme.

As the main contribution, we present a framework that can be used to jointly design precise P2P motion trajectories for uncertain systems in real-time based on the closed-form solutions provided, and to identify the flexible modes if any. In Section 2, the design of suppress-excite signals is presented, followed by outlining the flexible mode identification method in Section 3. The simulation results are given in Section 4, and final remarks are given in Section 5.

2. On-line point-to-point trajectory generation and embedded input shaping

A Visual abstract highlighting the main contributions of the current investigation, and their corresponding subsections, is give in Fig. 2. Firstly, the desired frequency range is split into three types of frequency bands (motion-related suppress band, detection-related excite band, and do-nothing band). Secondly, the needed time and frequency templates are designed using function composition, tuned using suitable optimization, and combined using suitable transformation while off-line. Finally, the overall

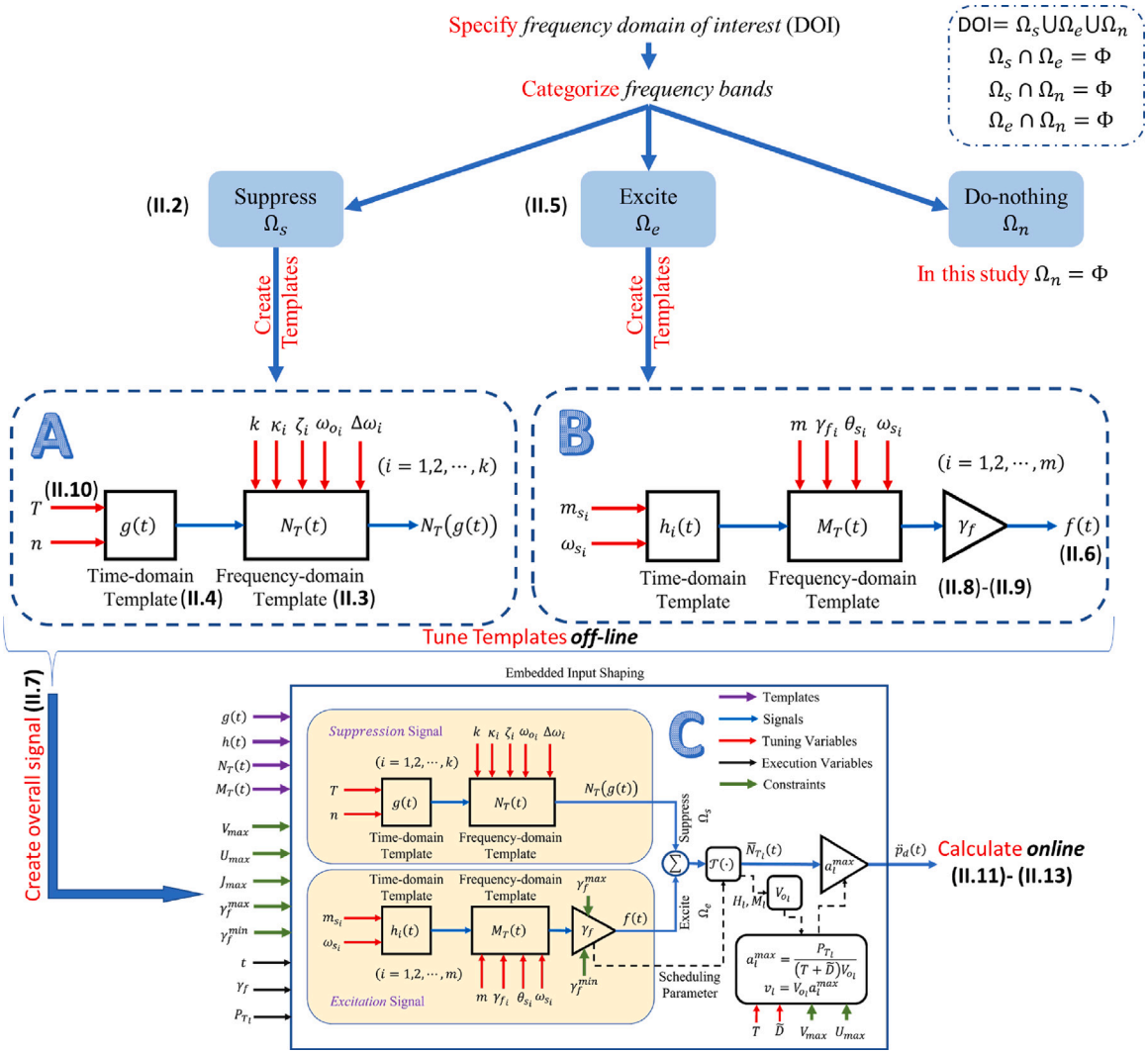

Fig. 2. Visual abstract highlighting the main contributions of the current investigation, and their corresponding subsections.

suppress-excite signal is used to generate online point-to-point trajectories to steer the motion system. Fig. 2.A shows the interaction of the time-domain and the frequency-domain templates used to generate the suppression signal of the acceleration pulse under embedded input shaping (EIS). Fig. 2.*B* shows the block diagram of the excitation signal used in this study. Fig. 2.C shows the generalized block diagram of the proposed combined suppress-excite method used in this study, $\forall t \in [0, T]$, according to embedded input shaping such that $\Omega_e \cap \Omega_s = \emptyset$, and $l = 1, 2, 3, 4$ denotes the used Case index. Note that J_{max} is used implicitly. In this section, we will be handling the following points:

- Describing the model that will serve as our benchmark is covered in Section 2.1. It is divided into two main parts: a modeled part including uncertainty, and an unmodeled part due to any available lightly-damped non-rigid body modes.
- To suppress these lightly-damped modes, we propose using a notch filter in the making of reference trajectories, where the filter tabs are matched to the frequency ranges of these modes. This will give rise to the so-called suppression signal. This is covered in Sections 2.2 through 2.4.
- To boost the tracking performance further, we introduce a set of excitation signals to facilitate the identification of the unmodeled lightly-damped modes before updating the suppression signal based on the detection information, i.e., adding the detected flexible modes to the list of frequencies to be suppressed. This is covered in Sections 2.5, 2.6, 2.8, 2.9, and 3.
- We opt to superimpose excitation signals onto the suppression signal, where their operating frequencies do not overlap. This is covered in Section 2.7. This ensures a detection-while-in-motion scheme of the reference trajectories. Such a feature is of a paramount importance for specific applications, e.g. machine health monitoring. After normalization, this suppress-excite signal will be used as the template when trajectory extrapolation is handled.

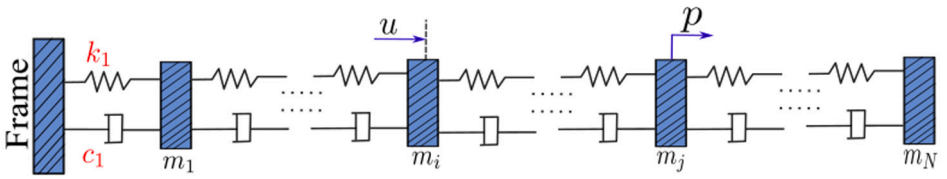

Fig. 3. A generic multi-mass-block noncollocated model inspired by [33,34].

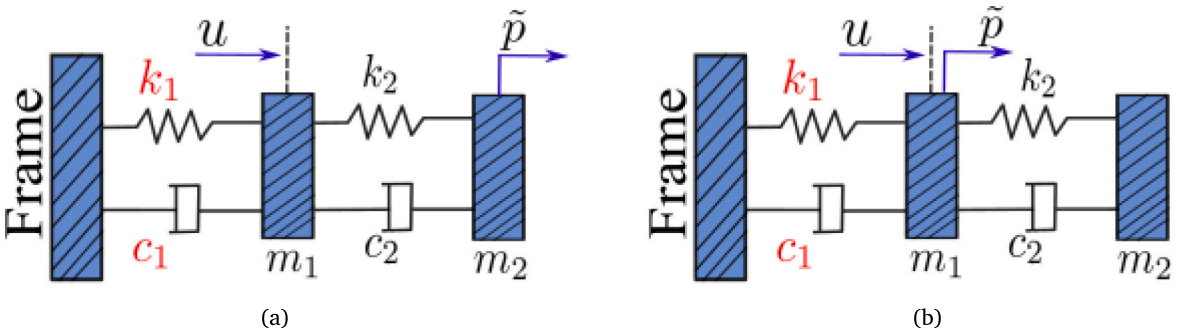

Fig. 4. Generic double-mass-block (a) noncollocated (b) collocated models.

- In Section 2.10, we investigate the relations between key tuning parameters to understand their effects on the designed reference trajectories.
- Imposing the available kinematical constraints on the designed trajectories is also discussed, along with the on-line extrapolation of the reference trajectories. This is covered in Section 2.11 through 2.13.
- Finally, for the sake of completion, we highlight how polynomial-based trajectory generation methods can benefit from the herein proposed approach as explained in Section 2.14.

2.1. Model description

According to [33,34], a precise and generic model of a friction-free ultra-precision system in a single direction of motion can be given by a multi-mass-block model depicted in Fig. 3 that consists of one *desired* rigid body mode and $N-1$ *undesired* non-rigid body (or flexible) modes as

$$G(s) = \frac{p(s)}{u(s)} = \frac{1}{m_t\, s^2} + \frac{1}{m_t} \sum_{i=2}^{N} \frac{\bar{\alpha}_i}{s^2 + 2\,\bar{\zeta}_i\, \bar{\omega}_i\, s + \bar{\omega}_i^2} \tag{1}$$

where $m_t = \sum_{i=1}^{N} m_i$ is the total mass, N is the number of distinct mass blocks – here used to denote the modes count –, both $\bar{\zeta}_i$ and $\bar{\omega}_i$ are the associated damping ratio and natural frequency of the ith mode, and $\bar{\alpha}_i$ is determined by the eigenmodes, the input u, and the output p of the system [33]. Oftentimes, the resonant mode associated frequencies are *unknown* a priori; and therefore need to be identified, or estimated [34].

In typical positioning applications [7,8,35], the *partially known* dynamics of the driven stage is usually approximated using a non/collocated double mass–spring-damper model as depicted in Fig. 4 with other variants still possible. When parametric uncertainties are involved, then the noncollocated model depicted in Fig. 4(a) is given as

$$G_p(s) = \frac{\tilde{p}(s)}{u(s)} = \frac{c_2\, s + k_2}{a_4\, s^4 + a_3\, s^3 + a_2\, s^2 + a_1\, s + a_0} \tag{2}$$

where $a_4 = m_1 m_2, a_3 = (m_1 + m_2)c_2 + m_2 c_1, a_2 = (m_1 + m_2)k_2 + m_2 k_1 + c_2 c_1, a_1 = (k_1 + k_2)c_1$ and $a_0 = k_1 k_2$ with m_1, m_2, c_1, c_2, k_1 and k_2 being uncertain parameters with values given as

$$\begin{aligned} m_{1,2} &= \widetilde{m}_{1,2}\,(1 \pm \Delta m_{1,2}) \\ c_{1,2} &= \widetilde{c}_{1,2}\,(1 \pm \Delta c_{1,2}) \\ k_{1,2} &= \widetilde{k}_{1,2}\,(1 \pm \Delta k_{1,2}) \end{aligned} \tag{3}$$

with $\widetilde{m}_{1,2}, \widetilde{c}_{1,2}$ and $\widetilde{k}_{1,2}$ representing the nominal values, and $\Delta m_{1,2}$, $\Delta c_{1,2}$, $\Delta k_{1,2}$ denote the corresponding *symmetric* relative uncertainties- *as percentages*. Recalling (1), the undesired flexible mode dynamics can be given as

$$G_{FM}(s) = \frac{\hat{p}(s)}{u(s)} = \sum_{i=3}^{N} \frac{\bar{\alpha}_i / m_t}{s^2 + 2\,\bar{\zeta}_i\,\bar{\omega}_i\,s + \bar{\omega}_i^2} \tag{4}$$

Combining both (2) and (4) according to (1) results in

$$\tilde{G}(s) = G_p(s) + G_{FM}(s) \tag{5}$$

and therefore the output of the system is the combined effect of all modes evaluated at the sensor location [34], i.e., $p(t) = \tilde{p}(t) + \hat{p}(t)$. Recalling (3), it is reasonable to consider the ith flexible mode parameters, i.e. $\bar{\alpha}_i, \bar{\zeta}_i, \bar{\omega}_i$, as uncertain even after being identified, or estimated. When $\bar{\alpha}_2 = -1$ [33], and both k_1 and c_1 representing the structural connection between the actuator block m_1 and the stage frame are ignored, i.e., identically zero, then $\tilde{G}(s) \equiv G(s)$. In this study, we treat $G_p(s)$ as the modeled uncertain part of the driven stage that produces the desired stage output $\tilde{p}(t)$, and $G_{FM}(s)$ as the unmodeled part corresponding to the flexible modes that produces the undesired stage output $\hat{p}(t)$ that may limit the used controller bandwidth. In this investigation, we use the system output $p(t)$ in the feedback control loop, and identify the undesired flexible mode frequencies without using state-observers [34]. Moreover, we keep k_1 and c_1 nonzero resembling a simplified model of the connection from the stage to the frame, e.g. via dynamic links or cable schlepp (carrier) that is used to transport cooling fluids and/or electrical signals from sensors and actuators. Ideally, it is desirable to have both k_1 and c_1 as zero such that the motion stage acts as a point-mass, see (1). However, this is impractical due to the previously mentioned reason, and therefore they are kept nonzero in this study. The values of both k_2 and c_2 can be larger/smaller than those of k_1 and c_1 depending on the designed motion stage and the actuators used. For example, a Lorentz actuator has a zero stiffness [3] and can be used to provide a natural decoupling of frame vibrations to the driven stage [2], and therefore k_2 is less than $k_1 > 0$. Similarly, the value of c_2 depends on the damping method in effect, e.g. active or passive damping [3]. In this study, $\{k_1, k_2\}$, and $\{c_1, c_2\}$ are treated as uncertain parameters with similar bounds and different nominal values as given in (3). Note that the same analysis still holds for generic cases where different bounds and nominal values of these parameters are used. A detailed analysis of the cable carrier effect on stage dynamics is available in [36–38].

Let the feedback controller used be given as

$$G_c(s) = k_p + k_i \frac{1}{s} + k_d \frac{s}{\tau_c\, s + 1} \tag{6}$$

which corresponds to a PID-based controller with k_p, k_i and k_d as its robustly-designed proportional, integral and derivative gains while τ_c is the controller (low-pass) filter time constant. Although more sophisticated variants of the PID controller are used in the literature [33,34], we found (6) to be sufficient according to the study objectives.

A widely used control strategy in positioning applications is a combined feedback-feedforward control scheme [7,8,35]. This control strategy helps in achieving perfect input-tracking if the *exact* plant inverse is used in the feedforward controller [33,39], and the feedback controller is used to reject errors due to non-zero initial conditions, measurement noise, and external perturbations [39].

Model-inversion design techniques of the feedforward controller [39–43] rely heavily on the available plant model. When the plant is uncertain, e.g. (2), then the *invertible* nominal plant model can be used instead to generate the model-inverse [39]. Since in this study we assume that the driven motion stage is both flexible and uncertain, we will use the nominal stage model of G_p to obtain the feedforward controller through model-inversion according to [39] while ignoring the *initially* unknown flexible mode dynamics, i.e., G_{FM}. Once these dynamics become known, then the resulting *parametric* flexible mode model can be used to update the feedforward controller as described in [33,40], if desired.

Remark 2.1. In this study, we aim at achieving better input-tracking performance while keeping the initially designed control loops intact. Therefore, the available model-based and data-based information about the uncertainties and the flexible modes – at any point in time – will be utilized in building or updating the dynamically friendly trajectories at the input side according to the herein proposed framework.

Recalling (2), the plant inverse using the nominal values ($\widetilde{G}_p^{-1}$)- where $\widetilde{G}_p^{-1}(s) \equiv G_p^{-1}(s)$ with $m_{1,2} = \widetilde{m}_{1,2}, c_{1,2} = \widetilde{c}_{1,2}$ and $k_{1,2} = \widetilde{k}_{1,2}$- is improper. As a remedy, a feedforward filter *preferably* given as

$$F(s) = \left(\frac{1}{\tau_f\, s + 1}\right)^{n_f} \tag{7}$$

is used in cascade with $\widetilde{G}_p^{-1}$ in the feedforward path, where $\tau_f > 0 \in \mathbb{R}$ is the time constant of the feedforward filter and $n_f \geq 3 \in \mathbb{Z}$ is chosen such that $F(s)\,\widetilde{G}_p^{-1}(s)$ is at least proper. Therefore, according to Fig. 5, the overall transfer function from $p_d(t)$ to $p(t)$- where the external disturbance $d(t)$ and the measurement noise $\nu(t)$ are assumed to be zero- is given as

$$G_{cl}(s) = \frac{p(s)}{p_d(s)} = \underbrace{\frac{\tilde{G}(s)\,G_c(s)}{1 + \tilde{G}(s)\,G_c(s)}}_{\text{Due to Feedback}} + \underbrace{\frac{\tilde{G}(s)\,F(s)\,\widetilde{G}_p^{-1}(s)}{1 + \tilde{G}(s)\,G_c(s)}}_{\text{Due to Feedforward}} \tag{8}$$

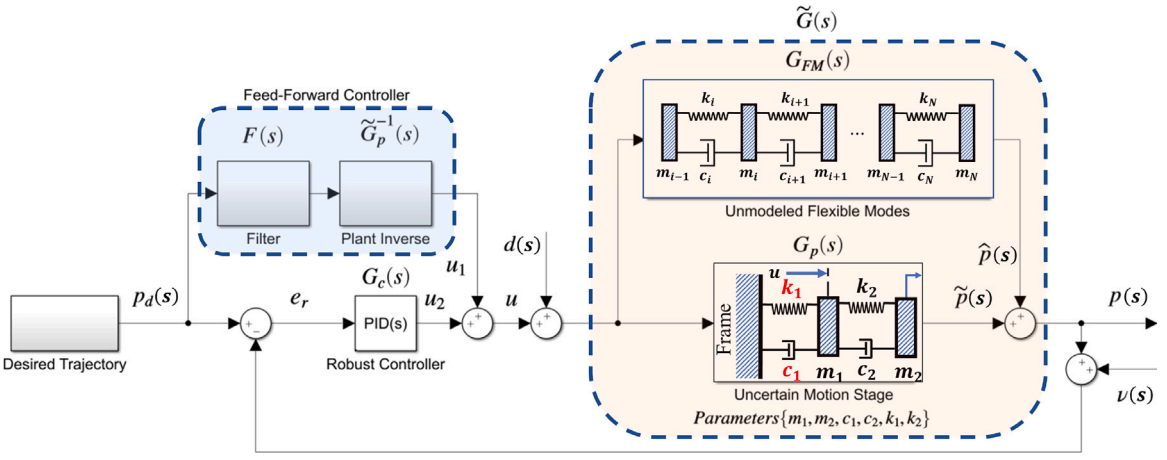

Fig. 5. The Simulink model of the uncertain flexible motion stage used in this study.

Remark 2.2. When a collocated model of the motion system is used – see Fig. 4(b) –, the plant G_p may have non/minimum phase lightly damped zeros [44,45]. Once inverted, the minimum-phase lightly damped zeros in G_p will contribute in resonance, and therefore, their corresponding frequencies can be added to the set of frequencies to be suppressed. Non/minimum phase lightly damped zeros will be handled in a separate study.

Based on (8), the lightly-damped oscillatory (flexible) modes originate due to some – or possibly all – poles in the feedback loop, driven stage $\tilde{G}(s) = G_p(s) + G_{FM}(s)$, and $F(s)\,\widetilde{G}_p^{-1}(s)$ transfer functions. Using the filter given by (7), the transfer function $F(s)\,\widetilde{G}_p^{-1}(s)$ does not have oscillatory frequencies, so we can focus on these frequencies associated with $\tilde{G}(s)$. Consequently, in this study, we first utilize the prior knowledge available about $G_p(s)$ to design a robust feedback controller $G_c(s)$ and a nominal feedforward controller $F(s)\,\widetilde{G}_p^{-1}$ while assuming (for the time being) $G_{FM}(s) = 0$. When the resulting feedforward–feedback control scheme from $p_d(t)$ to $\tilde{p}(t)$ is deemed *sufficient* to drive the uncertain stage- which can be tested for example by simulation or initial experiment-, we then design a special *suppress* acceleration signal based on the available lightly-damped oscillatory modes- excluding those of $G_{FM}(s)$ which still need to be identified. Later, to refine the tracking performance of the motion stage, an excitation signal will be added to the designed suppression signal in an attempt to identify the stage flexible modes if any, i.e., for the case when $G_{FM}(s) \neq 0$.

Remark 2.3. In general, external disturbance $d(t)$ and measurement noise $\nu(t)$ are handled by the feedback controller; however, in this study we assume both to be zero. Handling modeled $d(t)$ under the proposed framework will be investigated in future work.

Recall the noncollocated model, the collocated double-block model depicted in Fig. 4(b) can be handled similarly *assuming the nominal plant to be minimum phase*.

2.2. Uncertain lightly-damped modes

Based on the available prior knowledge, first principles modeling, estimation, and according to Fig. 5, the frequency bands associated with the uncertain lightly-damped modes can be obtained by using (2) through (7), then by taking sufficient samples of the uncertain parameters and observing the resulting variations of resonance frequencies while keeping (6) and (7) fixed.

Let $\Omega_s = \{\Omega_{s_1}, \ldots, \Omega_{s_k}\}$ be the set of all existing frequency bands associated with these modes, and $1 \leq k \in \mathbb{Z}$ is the number of bands. Additional undesired frequencies that need to be suppressed can be added to Ω_s, and may not be motion-related, e.g. electric resonance in motor drives [46,47], and nearby machine dynamics. Having attenuated frequency contents of the reference trajectory in Ω_s results in less excitation of the uncertain lightly-damped modes, and therefore better tracking is expected. The ith suppression band is given as

$$\begin{aligned}\Omega_{s_i} &= \{\omega \in \mathbb{R} \mid \omega_{o_i} - \Delta\omega_i \leq \omega \leq \omega_{o_i} + \Delta\omega_i\} \\ &= \{\omega \in \mathbb{R} \mid \omega_{s_i} \leq \omega \leq \omega_{e_{i+1}}\}\end{aligned} \tag{9}$$

where ω is the radial frequency (rad/s), ω_{o_i} is the band central frequency, $i = 1, 2, \ldots, k$, and $2\Delta\omega_i > 0$ denotes the rejection band with attenuation equals 3 dB. A generic example of Ω_s suppression bands is depicted in Fig. 6.

2.3. Cascaded notch filters

One way to remove the undesired frequency bands from the reference signal is to use filters [48,49] at the input side as input shapers (IS) [9] or within the closed-loop control system [23]. In this study, the overall impulse response of a k-cascaded notch

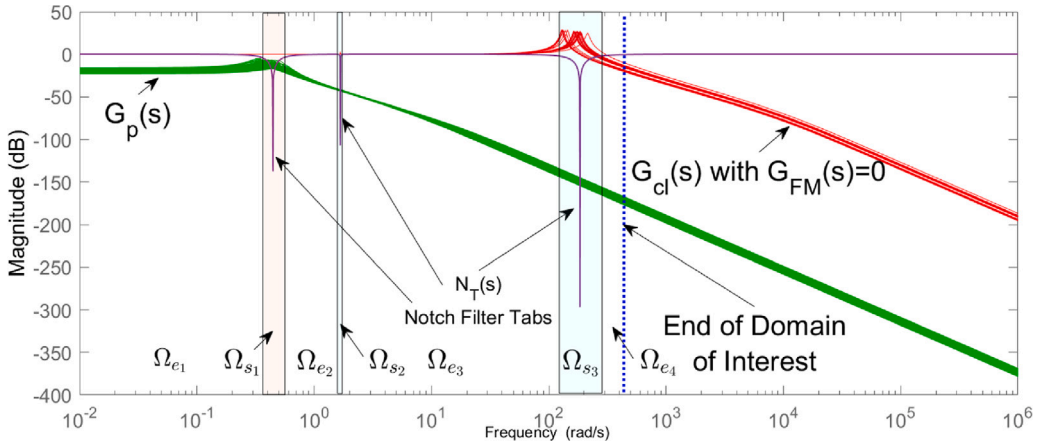

Fig. 6. A generic example that shows the known lightly-damped modes and their corresponding suppression bands Ω_{s_i} matched to the notch filter tabs. The excitation bands that both complement Ω_s in the domain of interest and may contain any unmodeled lightly-damped modes are denoted by Ω_{e_i}.

filter is used to construct the reference signal $p_d(t)$ while minimizing the undesired frequencies given in Ω_s. This is referred to as *embedded input shaping* (EIS).

The overall transfer function of the k-cascaded *distinct* notch filters under study is given as

$$N_T(s) = \prod_{i=1}^{k} \frac{s^2 + \omega_{o_i}^2}{s^2 + 2\,\kappa_i\,\zeta_i\,\omega_{o_i}\,s + \omega_{o_i}^2} = 1 + \sum_{i=1}^{k} \frac{\alpha_i^*}{s - \tau_i} + \frac{\alpha_i}{s - \tau_i^*} \tag{10}$$

where $0 < \kappa_i\,\zeta_i < 1$, the ith-filter scalar damping ratio $\zeta_i > 0 \in \mathbb{R}$ corresponding to the ith-frequency band, $\kappa_i > 0 \in \mathbb{R}$ is a tuning parameter, $(\cdot)^*$ denotes the complex conjugate, and the coefficients $\{\alpha_i, \tau_i\}$ are to be determined by the user, for example by using the *residue* MATLAB function [50], or by a lengthy process of partial fractions' decomposition. The quantity $2\,\kappa_i\,\zeta_i\omega_{o_i}$ determines the width of the rejected band that can be linked to the uncertainties of the driven stage through utilizing $2\Delta\omega_i$. Consequently, $\zeta_i \geq \zeta_i^{min} = \frac{\Delta\omega_i}{\kappa_i\,\omega_{o_i}}$ must be calculated based on each frequency band given in Ω_s. Ignoring the right most term in (10), it is clear that we *generally* have four cases based on the damping ratios, Case 1: $\zeta_i < 1/\kappa_i$, Case 2: $\zeta_i = 1/\kappa_i$, Case 3: $\zeta_i > 1/\kappa_i$ all $\forall i = 1, 2, \ldots, k$, and Case 4 has arbitrary values of $\forall\zeta_i > 0$. Case 4 establishes the link between the design process of cascaded IIR filters [51] and the proposed framework. In this study, we focus on Case 1 and Case 4 specifically for a mix of $\zeta_l < 1/\kappa_l$ and $\zeta_i > 1/\kappa_i$. Therefore, the filter impulse response for Case 1 using (10) is given as

$$N_T(t) \equiv N_{T_1}(t) = \delta(t) + \sum_{i=1}^{k} \alpha_i^*\,\mathrm{e}^{\tau_i\,t} + \alpha_i\,\mathrm{e}^{\tau_i^*\,t} \tag{11}$$

where $\tau_i = -\omega_{o_i}\left(\zeta_i - j\sqrt{1-\zeta_i^2}\right) = -\omega_{o_i}\,\zeta_i + j\,\omega_{d_i}$, and $\delta(t)$ is the *Dirac* delta function. Fig. 6 shows the matched notch filter tabs with the readily known lightly-damped modes bands.

The impulse response $N_T(t)$ of the used filter, i.e., here it relates to a notch filter, serves as the *frequency-domain template* that will assign various frequency-domain characteristics to the inputted time-domain scalar signal, e.g. $g(t) \in \mathbb{R}$ representing the raw signal, through function composition, i.e., $N_T(g(t))$.

2.4. Suppression chirps signals

Consider the following polynomial signal $\forall t \in [0, T]$ given as

$$g(t) = \left(\frac{T}{2}\right)^n - \left(t - \frac{T}{2}\right)^n \tag{12}$$

with even degree $n > 0 \in \mathbb{Z}$. The polynomial $g(t)$ serves as the *time-domain template* of the nonzero parts of the acceleration, the jerk, or other higher-order derivative trajectories. However, in this study, we focus mainly on designing the acceleration signal to generate S-curve-like profiles [52]. So, when $n = 2$, then the *snap is constant* [20,53], and this is the main motion profile to be studied herein. Note that having $n = 2$ in (12) also helps in writing the acceleration signal in a closed-form. Fig. 7 shows the normalized time-domain template $\tilde{g}(t) = g(t)/g(T/2)$ with $n = 2, 4, 6, 8, 10$ and $T = 0.08$ s as given in (12).

According to Newton's second law, the acceleration is directly proportional to the applied net force. Therefore, the applied force to the driven stage, i.e., $u(t)$ in Fig. 5, inherits the acceleration frequency contents of the reference signal, i.e., $g(t)$, shaped using embedded input shaping, i.e., EIS, to produce $\ddot{p}_d(t)$. Consequently, the excitation of the lightly-damped modes can be suppressed by

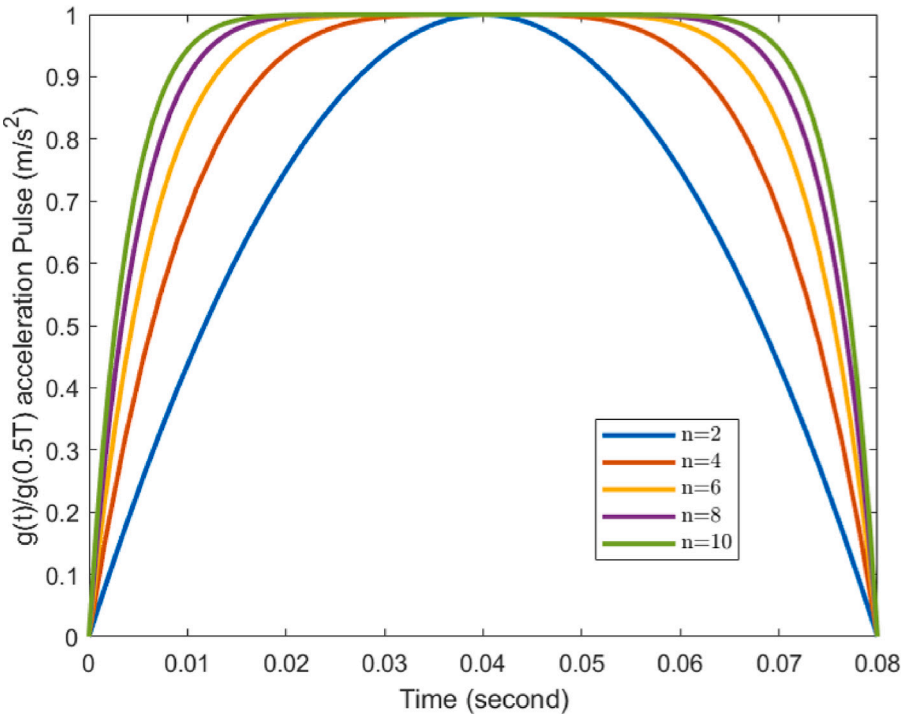

Fig. 7. The normalized time-domain template $\tilde{g}(t) = g(t)/g(T/2)$ with $n = \{2, 4, 6, 8, 10\}$, and $T = 0.08\,\mathrm{s}$ as given in (12).

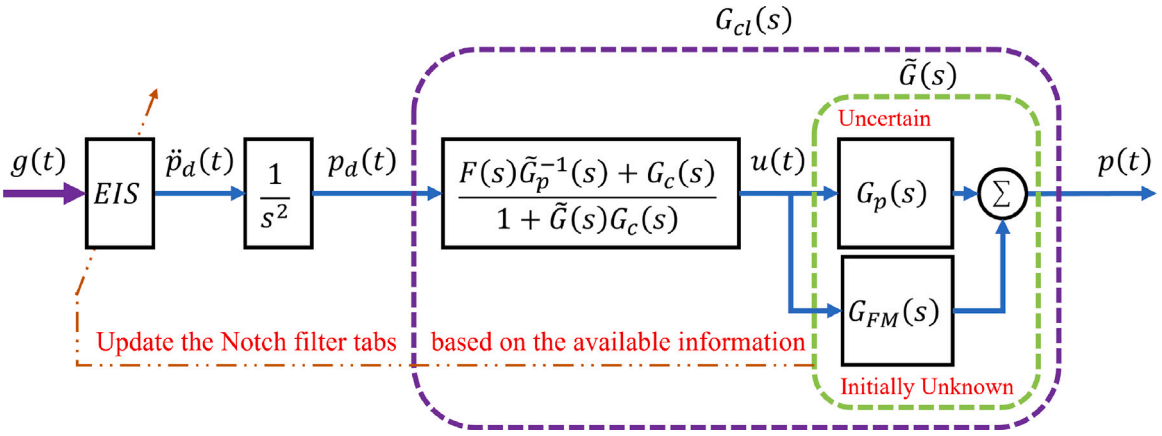

Fig. 8. The block diagram capturing the relations between the involved signals during the suppression mode of operation. There is more into the EIS block-depicted in Fig. 2.C- that will be explained shortly. To avoid stability issues, the EIS block update is done off-line.

crafting the frequency contents of the acceleration signal [5]. The interaction between the previously mentioned signals is captured in Fig. 8, which shows the relation between the needed force $u(t)$ and the EIS-crafted acceleration signal $\ddot{p}_d(t)$ given as

$$\begin{aligned} u(s) &= \left\{ \frac{F(s)\,\widetilde{G}_p^{-1}(s) + G_c(s)}{1 + \left(G_p(s) + G_{FM}(s)\right)\,G_c(s)} \right\} p_d(s) \\ &= M(s)\,p_d(s) = M(s)\,\frac{1}{s^2}\,\{\mathcal{L}\{\ddot{p}_d(t)\} + s\,p_d(0) + \dot{p}_d(0)\} \end{aligned} \tag{13}$$

where $\mathcal{L}\{\cdot\}$ denotes the Laplace transform, and $M(s)$ act as a frequency-dependent mass [54,55]. From (13), it is clear that $u(s)$ is affected by $G_{FM}(s)$ at any given time instant, when the control loops are kept unchanged, see Remark 2.1. Therefore,- as will be covered in a separate section- $\ddot{p}_d(t)$ should be *updated* to at least avoid exciting $G_{FM}(s)$ whenever new information is obtained about it, as suggested by Fig. 8. Also, note that when $\ddot{p}_d(t) = 0$, then $u(t) = 0$, and the forced vibration – if any – will vanish. Consequently, we may limit our interest in the interval where $\ddot{p}_d(t) \neq 0$, i.e., $\forall t \in [0, T]$.

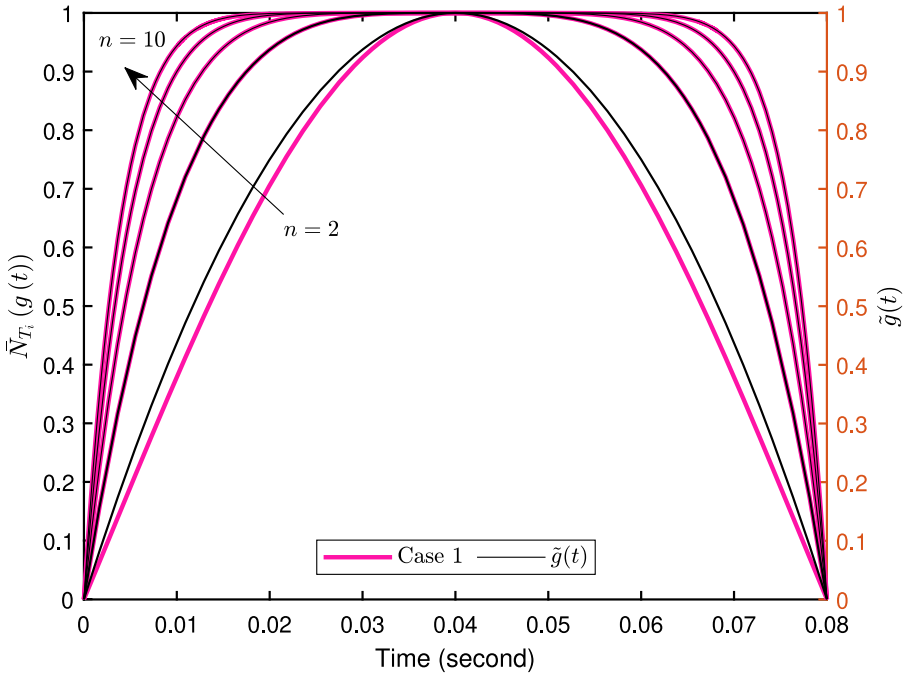

Fig. 9. A generic example that shows the shifted and scaled acceleration pulses $\bar{N}_{T_i}(g(t))$ under Case 1, resulting from applying EIS to $g(t)$ with $T = 0.08$ s, all being compared to the polynomial – without EIS – $\tilde{g}(t) = g(t)/g(T/2)$ with $n = \{2, 4, 6, 8, 10\}$ as given in (12).

Remark 2.4. When a detailed model that includes for example the viscoelastic behavior of the cable-schlepp is used, feedforward force is needed when the acceleration becomes zero because of the memory effect in the viscoelastic components [36–38]. This issue will be investigated in a separate study.

To realize the point-to-point motion, let us now create an *acceleration pulse* using EIS. The composition $N_{T_1}(g(t))$ generates the *suppression* signal of the acceleration pulse, which has in its construction preferred frequency and time domain contents [31]. The interaction between the two templates is depicted in Fig. 2.A. Mathematically, $\forall t \in [0, T]$, it is given as

$$\begin{aligned} N_{T_1}(g(t)) &= \delta(g(t)) + \sum_{i=1}^{k} \beta_i \, \mathrm{e}^{-\tau_i \left(t-\frac{T}{2}\right)^2} + \beta_i^* \, \mathrm{e}^{-\tau_i^* \left(t-\frac{T}{2}\right)^2} \\ &= \delta(g(t)) + \sum_{i=1}^{k} C_i(t) \, \sin\left(\omega_{d_i} \, g(t) - \Phi_i\right) \\ &= \delta(g(t)) + \sum_{i=1}^{k} C_i(t) \, \sin\left(\omega_{d_i} \, (T-t)t - \Phi_i\right) \end{aligned} \tag{14}$$

where $\beta_i = \alpha_i^* \, \mathrm{e}^{\tau_i \left(\frac{T}{2}\right)^2}$, $C_i(t) = 2|\alpha_i| \, \mathrm{e}^{-\sigma_1 \, g(t)}$, $\sigma_1 = \zeta_i \omega_{o_i}$, and $\Phi_i = tan^{-1}\left(\frac{\mathrm{Re}\{\alpha_i\}}{\mathrm{Im}\{\alpha_i\}}\right)$.

The signal $N_{T_1}(g(t)) - \delta(g(t))$ is the summation of k-chirp signals. Fig. 9 shows a comparison between the shifted and scaled acceleration pulses under Case 1, resulting from applying EIS to $g(t)$ with $T = 0.08$ s, and $\tilde{g}(t) = g(t)/g(T/2)$ with $n = 2$ as given in (12). The differences reveal how the composition affects the contents of $g(t)$ in order to avoid exciting the *known* lightly-damped modes. Compare this to Fig. 10 that shows the effect of applying convolution to the same previous cascaded notch filter and $g(t)$ with $T = 0.08$ s, and $n = \{2, 4, 6, 8, 10\}$. Note that at $t = T$, the resulting convoluted signal is nonzero; due to the delay introduced under the convolution process with a notch filter. That is why shaping the input using convolution is not recommended when notch filters are used [5,9].

The acceleration pulse is used to build the acceleration profile, and consequently the overall kinematical motion profiles depicted in Fig. 11. Clearly, the initial and terminal velocity, and acceleration values are zero. This kind of motion profiles is suitable for pick-and-place, stepping phase in wafer scanners, and similar point-to-point applications.

Remark 2.5. Interestingly, the time-domain effect of the EIS on $g(t)$ is obvious from (14), and Fig. 11 when $t = 0$ or $t = T$. By *trimming* the acceleration signal $N_T(g(t))$ such that $t \in (0, T)$, the contribution of the Dirac delta function appearing in (14) will be zero, and the jerk signal is discontinuous at the boundaries. Therefore, the snap will be infinite, which no real motor/actuator can provide. Compare this to the constant snap under (12) when $n = 2$. A finite snap can be recovered when working at the jerk level instead of the acceleration level.

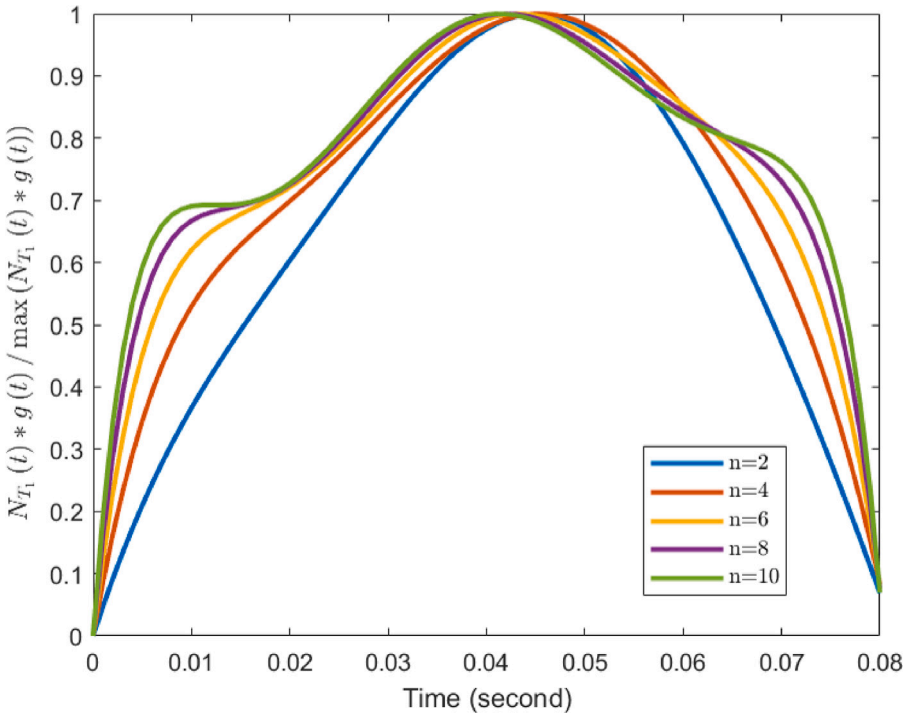

Fig. 10. A generic example that shows the shifted and scaled acceleration unit pulses under Case 1 resulting from applying convolution to a cascaded notch filter and $g(t)$ with $T = 0.08\,\mathrm{s}$, and $n = \{2, 4, 6, 8, 10\}$ as given in (12).

2.5. Excitation chirps signals

According to Fig. 8, the achieved position $p(t)$ is affected by the existing lightly-damped modes G_{FM}. Therefore, if the obtained tracking performance, e.g. $|e_r(t)|^2$, is not satisfactory, then the control loops have to be updated, or the input signal has to be redesigned. Recalling Remark 2.1, we opt to redesign the input signal based on any available knowledge. To gain such knowledge, an estimation [34], or detection [56] method must be involved, which obviously incorporates excitation signals. Some examples of the excitation signals are chirp, and (random) multisinusoidal signals [57].

In theory [58], a multisinusoidal signal leads to the best linear approximation if it is superimposed with a constant velocity, which will limit the excitation amplitude. Due to this limitation along with the low crest factor of the excitation, a multisinusoidal signal may lose its significance [58]. In general, one can state that the best type of excitation signals used is application driven [57,59]. In this study, we *choose* to utilize EIS to create linear chirp excitation signals that enable detecting unmodeled flexible modes if any, i.e., $G_{FM}(s) \neq 0$, outside Ω_s in the partially-known driven uncertain motion systems while in motion, see Fig. 6. Using linear chirp signals ensures the existence of closed-form solutions of the trajectory making problem. The design of the chirp excitation signals is presented in this subsection, while the detection method is handled in Section 3.

The chirp or swept sinusoidal signals are widely used in applications like synthetic sounds, mechanical vibration [60], radar and system identification [61,62]. Consider the overall transfer function of the parallel connection of m transfer functions given as

$$M_T(s) = \sum_{i=1}^{m} \frac{\gamma_{f_i}}{\omega_{s_i}} \frac{\omega_{s_i} cos(\theta_{s_i}) + s\ sin(\theta_{s_i})}{s^2 + \omega_{s_i}^2} \tag{15}$$

where $\gamma_{f_i} \in \mathbb{R}$ is the gain that reflects the importance of the excitation band associated with the frequency ω_{s_i}, and $\theta_{s_i} \in [-\pi, \pi]$ is its associated shift angle. Setting $\gamma_{f_i} = 0$ will disable this excitation band. Taking the inverse Laplace transform of (15) results in

$$M_T(t) = \sum_{i=1}^{m} \frac{\gamma_{f_i}}{\omega_{s_i}} \sin\left(\omega_{s_i} t + \theta_{s_i}\right) \tag{16}$$

Now, consider the following polynomial given as

$$h(t) = \left(\left(\frac{m_s}{\omega_s}\right)t + 1\right)t \tag{17}$$

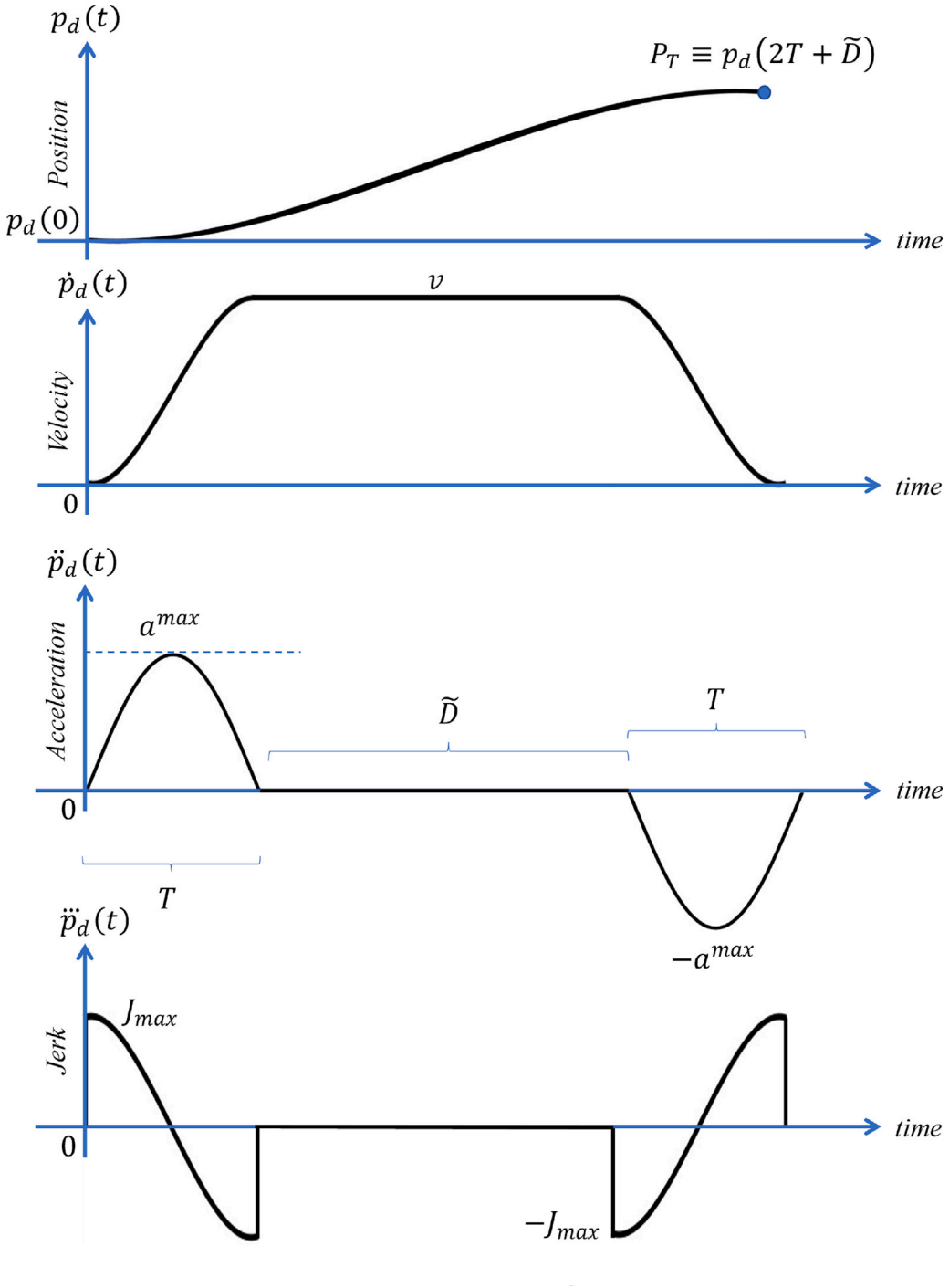

Fig. 11. The kinematical motion profiles under investigation.

and parametrized by m_s and ω_s. This polynomial, i.e., (17), will be used to describe the *linearly decreasing* frequency time-variation in the interval $\omega \in [\omega_e, \omega_s]$. A general form of the excitation signal $f(t)$ is given as

$$f(t) = \gamma_f \, M_T\left(h(t)\right) = \gamma_f \sum_{i=1}^{m} \frac{\gamma_{f_i}}{\omega_{s_i}} \sin\left(\omega_{s_i} \, h_i(t) + \theta_{s_i}\right)$$

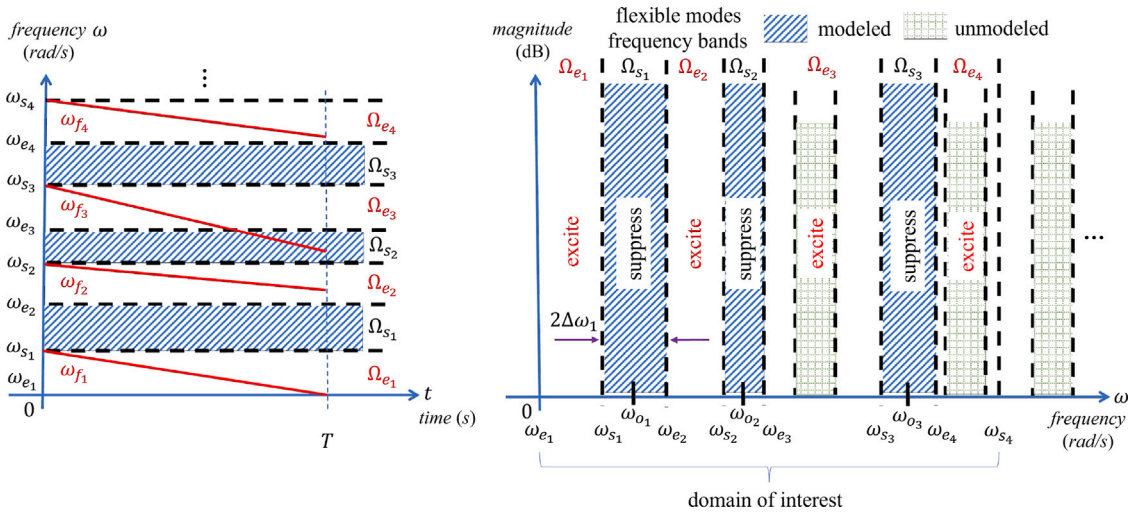

Fig. 12. The suppress-excite closed-loop frequency bands used in this study with linear variation of the excitation frequencies. This figure reveals how the simultaneous design of the suppress-excite signal is achieved using function composition. Interestingly, do-nothing bands can also be introduced to accommodate feedback controller dedicated notches, if desired.

$$= \gamma_f \sum_{i=1}^{m} \frac{\gamma_{f_i}}{\omega_{s_i}} \sin\left((m_{s_i} t + \omega_{s_i})t + \theta_{s_i}\right)$$
$$= \gamma_f \sum_{i=1}^{m} \frac{\gamma_{f_i}}{\omega_{s_i}} \sin\left(\omega_{f_i}(t)\, t + \theta_{s_i}\right) \tag{18}$$

where $\forall t \in [0, T]$, γ_f is the *global* excitation gain, and $\omega_{f_i}(t) \in \Omega_e = \bigcup_{i=1}^{m} \Omega_{e_i} = \bigcup_{i=1}^{m} [\omega_{e_i}, \omega_{s_i}]$ denoting the set of all frequencies to excite. Also, $\Omega_e \cap \Omega_s = \emptyset$. Comparing (14) and (18), it is clear that the latter is the summation of m-chirp signals with constant amplitudes. The generation of the excitation signal $f(t)$ is summarized in Fig. 2.B, and the interplay between Ω_e and Ω_s is exemplified in Fig. 12. After specifying Ω_s, the design engineer has to specify Ω_e and the frequency domain of interest (DOI) based on the application and the available information. Consequently, designing $f(t)$ can be commenced.

2.6. The design of the excitation signal $f(t)$

Using (12) ensures that the suppression signal is *symmetric*, which yields

$$N_{T_1}(g(0)) \equiv N_{T_1}(g(T)) \tag{19}$$

Now, to ensure the integration of the suppression and the *not necessarily symmetric* excitation signal – as explained shortly –, a uniform translation of $f(t)$ end-points, i.e., at $t = 0$ and $t = T$, must exist. Therefore, the boundary condition given by

$$f(0) \equiv f(T) \rightarrow \gamma_f \sum_{i=1}^{m} \frac{\gamma_{f_i}}{\omega_{s_i}} \sin\left(\theta_{s_i}\right) = \gamma_f \sum_{i=1}^{m} \frac{\gamma_{f_i}}{\omega_{s_i}} \sin\left(\omega_{f_i}(T)\, T + \theta_{s_i}\right) \tag{20}$$

must hold. Satisfying (20) can be done *off-line* either locally or globally, for each we provide a method by means of example.

- Method I (local design):

Taking the corresponding terms at both sides in (20), and utilizing the periodicity of the sine function results in

$$\sin(\theta_{s_i}) = \sin((m_{s_i} T + \omega_{s_i})T + \theta_{s_i}) \rightarrow,$$
$$(m_{s_i} T + \omega_{s_i})T = \pm 2\, L_i\, \pi \rightarrow m_{s_i} = \frac{\pm 2 L_i \pi - \omega_{s_i} T}{T^2} < 0 \tag{21}$$

with $L_i \geq 0 \in \mathbb{Z} \subset \left[\frac{\omega_{e_i} T}{2\pi}, \frac{\omega_{s_i} T}{2\pi}\right]$, otherwise either Ω_{e_i} cannot be swept entirely, or undesired frequency contents may appear such that $\Omega_e \cap \Omega_s \neq \emptyset$, see Fig. 12. In such a case, a global solution must be sought.

• Method II (global design):

The parameters $\chi_i = \{\gamma_{f_i}, m_{s_i}, \theta_{s_i}\}, \forall i = 1, 2, \dots, m$ can be used to solve (20) globally using the convex optimization problem – independent of γ_f – given as

$$\begin{aligned} \min_{\chi_i \in \chi} \quad & \mathcal{J}_e \\ \text{s.t.} \quad & -\omega_{s_i} D_i \le \gamma_{f_i} \le \omega_{s_i} D_i, \\ & \frac{\omega_{e_i} - \omega_{s_i}}{T} \le m_{s_i} < 0, \\ & -\pi \le \theta_{s_i} \le \pi, \\ & M_T(h(0)) = M_T(h(T)) \end{aligned} \tag{22}$$

where $D_i > 0 \in \mathbb{R}$ could be related to the signal-to-noise ratio inside Ω_{e_i}, and the performance index $\mathcal{J}_e$ is given as

$$\mathcal{J}_e = r_0\Big(\sum_{i=1}^{m} \frac{\gamma_{f_i}}{\omega_{s_i}} \sin\left(\theta_{s_i}\right)\Big)^2 + \sum_{i=1}^{m} r_i\left(\omega_{f_i}(T)\,T - \omega_{e_i}\right)^2 + r_0\Big(\sum_{i=1}^{m} \frac{\gamma_{f_i}}{\omega_{s_i}} \sin\left(\omega_{f_i}(T)\,T + \theta_{s_i}\right)\Big)^2 \tag{23}$$

with the weighting factors $r_l \ge 0 \in \mathbb{R}, \forall l = 0, 1, \dots, m$. The performance index (23) allows for minimizing the excitation signal contribution at the end-points, and at the same time allows for maximizing the frequency sweep in the excitation bands.

2.7. Building the unit-pulse suppress-excite signal

According to the proposed framework, the suppress, and suppress-excite are the two modes of operation used. In the former, point-to-point motion takes place while suppressing the readily known flexible modes in Ω_s. Similarly, in the latter mode of operation, point-to-point motion takes place while suppressing the readily known flexible modes in Ω_s, as well as exciting all the frequencies in Ω_e to detect any unknown flexible modes. Adopting an input–output perspective when dealing with the driven stages, we superimpose the suppression and excitation signals at the acceleration level – see (13) –, and let γ_f control the mode of operation. Fig. 2.C shows the superposition of the two signals, where *in general* time-domain template, frequency-domain template, and the tuning variables- *except* $\widetilde{D}$- are only used when a major update is needed, while the execution variables and $\widetilde{D}$ are always used in any operation mode. A major update is needed when new flexible modes are identified, and/or performance enhancement is required.

Remark 2.6. Based on the application, $\widetilde{D}$ plays a significant role in building the trajectory profiles. For example, in wafer scanners [2,3], the (almost) repetitive motion implies that P_T is fixed and known, i.e., the step and the scan lengths, and therefore $\widetilde{D}$ should be optimized and kept fixed, see [18]. When P_T is varying, then $\widetilde{D}$ must vary accordingly; to meet the kinematical constraints, as will be explained shortly.

Recalling the operating principle of notch filters, the available frequency contents of the suppression signal during the suppress mode of operation can be sufficient to detect the *unknown* flexible modes *outside* Ω_s. If this is not the case, then dedicated excitation signals are needed. So, adding $N_{T_1}(g(t))$ and $f(t)$ and applying the transformation ($\mathcal{T}$) yields

$$\bar{N}_{T_1}(t) = \begin{cases} 0 & \text{if } t \in (-\infty, 0] \\ \frac{1}{M_1}\left(N_{T_1}(g(t)) + f(t) - H_1\right) & \text{if } t \in (0, T) \\ 0 & \text{if } t \in [T, \infty) \end{cases} \tag{24}$$

where H_1 and M_1 are the DC-offset and the amplitude scaling factor given as

$$\begin{aligned} & H_1 = \lim_{t \to 0^+} W(t) \overset{(12),(20)}{\equiv} \lim_{t \to T^-} W(t) \\ & M_1 := \max\left\{ \left| W\left(\tilde{t}_i\right) - H_1 \right| : \left. \frac{dW}{dt} \right|_{t=\tilde{t}_i} = 0 \right\} \qquad \tilde{t}_i \in (0, T), \forall i \ge 1 \in \mathrm{Z} \end{aligned} \tag{25}$$

where $W(t) = N_{T_1}(g(t)) + f(t)$. Note that the existence of the uniform translation H_1 depends on the solution of (21) and (22). Interestingly, the *findpeaks* MATLAB function can be used to find M_1 for any given value of γ_f.

2.8. Relation between γ_f and the transformation $\mathcal{T}$

As suggested by Fig. 2.C and while fixing $M_T(h(t))$, it is possible to have only the suppression signal by setting $\gamma_f = 0$, or having the suppression signal along with different levels of excitation – to facilitate detecting the unmodeled flexible modes – by varying $0 < \gamma_f \le \gamma_f^{max}$. However, by carefully considering (25), it is clear that both M_1 and H_1 are functions of the excitation signal's global gain γ_f. Therefore, from a practical point of view, the transformation $\mathcal{T}$ must be scheduled based on γ_f, i.e., $\mathcal{T}(\gamma_f)$, preferably in a pointwise fashion for $0 < \gamma_f^{min} \le \gamma_f \le \gamma_f^{max}$, in any continuous space after applying the parametric constraints.

2.9. Choosing γ_f

The main idea behind using the suppress-excite signal with uncertain positioning systems is to concurrently *detect-while-in-motion*. Therefore, limiting γ_f below an upper bound γ_f^{max} helps in maintaining tracking performance, while having $0 < \gamma_f^{min} \leq \gamma_f$ helps in detecting the unmodeled flexible modes, in general. Recalling Fig. 2.C, (24) and under known symmetric jerk bounds ($\pm J_{max}$), the resulting jerk signal of the *reference trajectory* ($J(t) \equiv \dddot{p}_d(t)$) depicted in Fig. 11 must satisfy $\forall t \in [0, T]$ the condition given as

$$J_{max} \geq |J(t)| = \left| \frac{d\ddot{p}_d(t)}{dt} \right| = U_{max} \left| \frac{d\bar{N}_{T_1}(t)}{dt} \right| \rightarrow \left| \frac{d\bar{N}_{T_1}(t)}{dt} \right| \leq \frac{J_{max}}{U_{max}} \tag{26}$$

where $U_{max} > 0$ is the ultimate permissible acceleration value, and therefore, γ_f^{max} can be found. In this study, (26) is used to find γ_f^{max} by trial-and-error when T *is fixed*.

Finding γ_f^{min} is more difficult compared to finding γ_f^{max}, because the former is affected by the sensitivity of the identification/detection method used, and to a very extent the closed-loop dynamics of the uncertain stage and its flexible modes. Therefore, it seems reasonable to set $\gamma_f = \gamma_f^{max}$ when suppress-excite mode, i.e., $\gamma_f > 0$, is *occasionally* used, or when the tracking performance can be sacrificed in favor of flexible modes detectability.

2.10. Choosing T

Choosing the appropriate T that *suits the uncertain dynamics of the driven stage* is done either *off-line or on-line* for any choice of $\kappa_1 > 0, \zeta_i \geq \zeta_i^{min}, \forall i = 1, 2, \ldots, k$ through minimizing

$$\min_T \mathcal{J}_1 = \sum_{i=1}^{k} \gamma_i \, Q\left(\omega_{o_i}\right) \tag{27}$$

where $Q(\omega) = |\mathcal{F}\{\bar{N}_{T_1}(t)\}|^2$, the weights $\gamma_i \in \mathbb{R}, \forall i$, and the *truncated* Fourier transform [63] of (24) is given as

$$\begin{aligned} \mathcal{F}\{\bar{N}_{T_1}(t)\} = & \frac{-1}{M_1} \left\{ H_1 \, \delta(\omega) + \cdots j \sum_{i=1}^{k} \frac{\beta_i}{2} \sqrt{\frac{\pi}{\tau_i}} \exp\left(\sigma_2\right) \left\{ \operatorname{erf}\left(\frac{\omega + j\,T\tau_i}{2\sqrt{\tau_i}} \right) - \operatorname{erf}\left(\frac{\omega - j\,T\tau_i}{2\sqrt{\tau_i}} \right) \right\} \right. \\ & \left. + \frac{\beta_i^*}{2} \sqrt{\frac{\pi}{\tau_i^*}} \exp\left(\sigma_3\right) \left\{ \operatorname{erf}\left(\frac{\omega + j\,T\tau_i^*}{2\sqrt{\tau_i^*}} \right) - \operatorname{erf}\left(\frac{\omega - j\,T\tau_i^*}{2\sqrt{\tau_i^*}} \right) \right\} - \gamma_f \, \mathcal{F}\{M(h(t))\} \right\} \end{aligned} \tag{28}$$

where $\sigma_2 = \left(\frac{\omega^2}{4\tau_i} - j\frac{T\omega}{2} \right)$, $\sigma_3 = \left(\frac{\omega^2}{4\tau_i^*} - j\frac{T\omega}{2} \right)$, erf denotes the error function, and the Fourier transform of $M(h(t))$ is given as

$$\begin{aligned} \mathcal{F}\{M(h(t))\} = & \sum_{l=1}^{m} \eta_l \exp\left(\sigma_4\right) \left\{ \operatorname{erf}\left(\frac{\omega + j\omega_{s_l}}{2\sqrt{m_{s_l}}} \right) - \operatorname{erf}\left(\frac{\omega + j(\omega_{s_l} + 2T\, m_{s_l})}{2\sqrt{m_{s_l}}} \right) \right\} \\ & - \bar{\eta}_l \exp\left(\sigma_5\right) \left\{ \operatorname{erf}\left(\frac{\omega - j\omega_{s_l}}{2\sqrt{-m_{s_l}}} \right) - \operatorname{erf}\left(\frac{\omega - j(\omega_{s_l} + 2T\, m_{s_l})}{2\sqrt{-m_{s_l}}} \right) \right\} \end{aligned} \tag{29}$$

where

$$\eta_l = \frac{\gamma_{f_l}}{4\omega_{s_l}} \sqrt{\frac{\pi}{m_{s_l}}}, \qquad \bar{\eta}_l = \frac{\gamma_{f_l}}{4\omega_{s_l}} \sqrt{\frac{\pi}{-m_{s_l}}}, \quad \sigma_4 = \theta_{s_l} + \frac{\left(\omega + j\omega_{s_l}\right)^2}{4\, m_{s_l}} \quad , \sigma_5 = -\theta_{s_l} - \frac{\left(\omega - j\omega_{s_l}\right)^2}{4\, m_{s_l}}$$

2.11. Case 4 with $\zeta_l < 1/\kappa_l$ and $\zeta_i > 1/\kappa_i$

When the ith mode is non-oscillatory, i.e., $\zeta_i > 1/\kappa_i$, then the eigenvalues are real and distinct, i.e., the mode is over-damped. The same previous equations can be used by simply replacing τ_i^* with $\bar{\tau}_i$ to denote the second real and distinct eigenvalue of the ith mode. The associated coefficients are updated accordingly.

2.12. Calculating velocity profile

In positioning applications like lithography, machining, marking and laser cutting, special operations take place under constant velocity motion profiles denoted here by (V_{o_1}) and last for a specific time interval denoted here by ($\widetilde{D} \geq 0$). To stop the motion, a decelerating pulse is introduced at the end of the current motion profile. If both the accelerating and decelerating pulses are similar except for their signs, then the profile is deemed *symmetric*, otherwise it is *asymmetric*, see Fig. 11. Both types are attainable using the proposed framework. Integrating (24) yields

$$V_{o_1} = \int_0^T dt\, \bar{N}_{T_1}(t) = \frac{1}{M_1} \int_{\rightarrow 0}^{\rightarrow T} dt \left(N_{T_1}(g(t)) - H_1 \right) = \frac{1}{M_1} \left\{ \tilde{V}_{o_1} - H_1 \, T \right\}$$

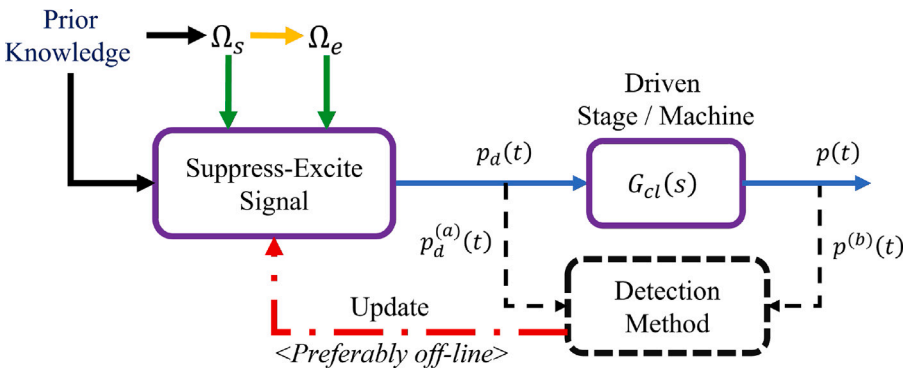

Fig. 13. Block diagram of the combined flexible modes detection and suppress-excite signal generation where the ath and bth time derivatives of the input and output signals, respectively, are used for detection. Different dashed arrows styles reflect various time update rates. To avoid stability issues, the needed update of the suppress-excite signal is done off-line.

$$\tilde{V}_{o_1} = \sum_{i=1}^{k} \left\{ \beta_i \sqrt{\frac{\pi}{-\tau_i}}\, \mathrm{erfi}\left(\frac{T\sqrt{-\tau_i}}{2}\right) + \beta_i^* \sqrt{\frac{\pi}{-\tau_i^*}}\, \mathrm{erfi}\left(\frac{T\sqrt{-\tau_i^*}}{2}\right) \right\} + \gamma_f \sum_{l=1}^{m} 2\sqrt{2}\eta_l \Big\{ \big(S(\sigma_6) - S(\sigma_7)\big)\cos(\sigma_8) + \big(C(\sigma_6) - C(\sigma_7)\big)\sin(\sigma_8) \Big\} \tag{30}$$

where erfi is the complex error function, S and C denote the Fresnel sine and cosine functions [64], and

$$\sigma_6 = \sqrt{\frac{2}{\pi\, m_{s_l}}}\Big(0.5\,\omega_{s_l} + T\, m_{s_l}\Big), \quad \sigma_7 = \frac{\omega_{s_l}}{\sqrt{2\pi\, m_{s_l}}}, \quad \sigma_8 = \frac{4\theta_{s_l} m_{s_l} - \omega_{s_l}^2}{4m_{s_l}}.$$

which makes V_{o_1} dependent on γ_f. The scanning or constant velocity after the acceleration pulse (v_1) is given as

$$v_1 = V_{o_1}\, a_1^{max} \tag{31}$$

with $|v_1| \le |V_{o_1}|\, U_{max} \le V_{max}$, where V_{max} is the maximum allowed speed, and the P2P *pitch size* (P_{T_1}) achieved using the EIS is given as

$$P_{T_1} = \left(T + \widetilde{D}\right) v_1 \tag{32}$$

Consequently, from (31) and (32), we may write

$$a_1^{max} = \frac{P_{T_1}}{\left(T + \widetilde{D}\right) V_{o_1}} \le U_{max} \tag{33}$$

2.13. Extrapolation of trajectories

Once T is determined – preferably using (27) – and the other tuning variables are kept fixed in a way that best suits the driven stage dynamics, the displacement P_{T_1} in (32) can be achieved in real-time by choosing both $\widetilde{D} \ge 0$ and a_1^{max} appropriately such that $|v_1| \le V_{max}$ and $|a_1^{max}| \le U_{max}$. Recalling Fig. 2.C and (33), it is clear that the tuning variable $\widetilde{D}$ can be used to determine a_1^{max} needed for any P_T displacement. For the applications that involve a task during the constant speed phase of the motion profile where $\widetilde{D} \gg T$, the condition given as $0 \le \widetilde{D}_{min} \le \widetilde{D} \le \widetilde{D}_{max}$ may be applied. Increasing $\widetilde{D}$ results in a reduced a_1^{max}, and therefore less tracking error, and reduced throughput are expected. The execution time of any point-to-point motion equals $2T + \widetilde{D}\, seconds$.

2.14. Relation with polynomial-based methods

According to the Weierstrass approximation theorem [65], every continuous function defined on a closed interval can be uniformly approximated by a polynomial function to any degree of accuracy. Because computers can easily evaluate polynomial functions in real-time, it is possible to approximate and replace (24) using polynomials of any desired order; so that polynomial-based methods for trajectory generation can benefit from it [5]. The costs associated with such approximations will be studied separately.

3. Flexible modes detection

Recalling Fig. 5, the problem formulation suggests that a closed-loop system identification (CLSI) should be involved. In this investigation, however, we are more interested in designing the suppress-excite signal for uncertain positioning applications with unmodeled flexible modes than deeply exploring the interplay between the designed signals and CLSI. Nevertheless, to show the applicability of the proposed suppress-excite signal design process in CLSI, we resort to use a weak form of identification, i.e., detection.

3.1. Detection method

Adopting a gray-box identification model, Fig. 13 shows the relation between the designed suppress-excite trajectory $p_d(t)$ and the closed-loop system G_{cl} response $p(t)$. The system G_{cl} given in (8) includes all control loops as well as the modeled and unmodeled flexible modes. When a lightly-damped flexible mode is excited, it will resonate and manifests itself as a peak in the system response when observed in frequency domain. A *generic* transfer function measurement is given as

$$\mathrm{M}_a^b(\omega) = \left| \frac{\mathrm{DFT}(p^{(b)}(t))}{\mathrm{DFT}(p_d^{(a)}(t))} \right| \tag{34}$$

where $\mathrm{DFT}(\cdot)$ denotes the (discrete) Fourier transform of the ath and bth time derivatives of the input $p_d(t)$ and the output $p(t)$, respectively. When $a = b$, then M_a^b denotes the transmissibility (TM). Let $a = 2$ and let the input $p_d^{(2)}(t)$ be associated with a *virtual* force, therefore when $b = \{0, 1, 2\}$, then M_a^b relates to the compliance, mobility and accelerance, respectively.

In this study, we use M_2^2 to detect the resonance frequencies of the unmodeled flexible modes using the designed suppress-excite signals. Note that detecting the damping ratios and mode shapes associated with the flexible modes is not necessary under the current suppress-excite signal design formulation; because we only need to know the flexible mode frequency to locate the suppress notch. Recalling Fig. 13, in time domain, the input $p_d^{(a)}(t)$ and output $p^{(b)}(t)$ measurements are obtained using appropriate sensors or estimators, where the ath and bth time derivatives are chosen based on the application.

Recalling Remark 2.1, the proposed technique will still be applicable when the involved controllers are digitally synthesized due to the adopted input–output perspective of the system. The only information needed is the frequencies of the flexible modes which can be given in s-domain representation using suitable transformations.

3.2. Utilizing detected flexible modes

To utilize the detected flexible mode frequencies, either a complete or partial update of the existing notch filters tunable parameters is carried out mainly by changing their damping ratios, introducing additional notch filters in cascade, i.e., increase k, or both. Note that this update will also affect the depths of the notches. The choice made depends on the number of the detected flexible modes and the distribution of their corresponding frequencies where a depiction of $\mathrm{M}_a^b(\omega)$ or $\bar{\mathrm{M}}_a^b(\omega)$ versus frequency will be insightful. This update process of the designed suppress-excite signal is highlighted by the red dashed arrow depicted in Fig. 13, which shows the flow of the design process of the suppress-excite signals. To avoid stability issues, we opt to update the involved parameters off-line.

According to Figs. 5 and 13, the tracking error $e_r(t)$- from input–output perspective- can be given as $e_r(t) = p_d(t) - p(t)$. Recalling Fig. 2.C, using (24) with $\gamma_f = 0$, and (34) with $a = b = 2$ yields

$$\begin{aligned} |\,\mathrm{FFT}\{e_r^{(2)}(t)\}|^2 &= |\,\mathrm{FFT}\{p_d^{(2)}(t)\} - \mathrm{FFT}\{p^{(2)}(t)\}|^2 \\ &= |1 - \mathrm{M}_2^2(\omega)|^2\, |\,\mathrm{FFT}\{p_d^{(2)}(t)\}|^2 \\ &= |a_{max}|^2\, |1 - \mathrm{M}_2^2(\omega)|^2\, |\,\mathrm{FFT}\{\bar{N}_{T_1}(t)\}|^2 \end{aligned} \tag{35}$$

where $t \in [0, T]$. Therefore, the tracking error can be minimized in real-time through minimizing $|a_{max}|$ which will then increase the execution time – see Section 2.13 – and therefore is not recommended, and/or off-line through minimizing the quantities $|\,\mathrm{FFT}\{\bar{N}_{T_1}(t)\}|^2$, and $|1 - \mathrm{M}_2^2(\omega)|$ mainly at certain observable frequencies of interest $(\forall \omega_l \in \Omega_o)$ with $|\Omega_o| = \bar{k}, l = 1, 2, \ldots, \bar{k}$. This is the essence of (27) for known $\zeta_i, \omega_{o_i}, i = 1, 2, \ldots, k$ parameters. Note that $\Omega_s \subseteq \Omega_o$, i.e., not every observable frequency has to have an associated tab in the k-cascaded notch filter. The elements of Ω_o can be obtained from the extreme points of $\mathrm{M}_a^b(\omega)$ resulting from the application of the designed suppress-excite signal. Minimizing $|1 - \mathrm{M}_2^2(\omega)|$ can be done by redesigning the feedforward control loop, e.g. as described in [33], or the feedback control loop. Recalling Remark 2.1, the minimization of $|\,\mathrm{FFT}\{e_r^{(2)}(t)\}|$ is achieved by minimizing $|\,\mathrm{FFT}\{\bar{N}_{T_1}(t)\}|$. Consequently, a generalization of (27) is given as

$$
\begin{aligned}
\min_{\theta \in \Theta} \quad & s_1 \mathcal{J}_s + s_2 U(\gamma_f)\, \mathcal{J}_e \\
\text{s.t.} \quad & 0 \le \omega_{o_i} \le DOI, \\
& 0 < \omega_{o_1} < \omega_{o_2} \cdots < \omega_{o_k}, \\
& T > 0,\ n = 2,\ \kappa_i > 0,\ \Delta\omega_i > 0, \\
& \frac{\Delta\omega_i}{\kappa_i\, \omega_{o_i}} \le \zeta_i, \\
& 1 \le k \le \bar{k}, \\
& 0 \le \gamma_f, \\
& \omega_{s_i} = \omega_{o_i} - \Delta\omega_i, \\
& \omega_{s_{k+1}} = DOI, \\
-\omega_{s_i} D_i \le \gamma_{f_i} \le \omega_{s_i} D_i,\ & D_i > 0, \\
& 0 \le \omega_{e_i} < \omega_{s_i}, \\
\frac{\omega_{e_i} - \omega_{s_i}}{T} \le m_{s_i} < 0, & \\
-\pi \le \theta_{s_i} \le \pi, & \\
M_T(h(0)) = M_T(h(T)) &
\end{aligned}
\tag{36}
$$

where $U(\cdot)$, is the unit-step function, $s_1, s_2 \ge 0 \in \mathbb{R}$ are weighting factors, the excitation index $\mathcal{J}_e$ is given in (23), the suppression index $\mathcal{J}_s = \sum_{l=1}^{\bar{k}} \gamma_l\, Q(\omega_l), \gamma_l \ge 0$, $Q(\omega) = |\mathcal{F}\{\bar{N}_{T_1}(t)\}|^2$- see (28)-, $i = 1, 2, \ldots, k$, and the set of decision variables $\Theta = \{T, \widetilde{D}, n, k, \kappa_i, \zeta_i, \omega_{o_i}, \Delta\omega_i, \gamma_f, m_{s_i}, \omega_{s_i}, m, \gamma_{f_i}, \theta_{s_i}\}$.

In (36), it is worth noting that γ_f^{max} is not specified; because we allowed T to vary. Also, when $\Delta\omega_i$ is associated with a detected flexible mode in G_{FM}, then it is treated as a decision variable, i.e., $\Delta\omega_i \in \theta$, to minimize the effect of biased detection. If $\Delta\omega_i$ is associated with a modeled flexible mode in G_p, it is not treated as a decision variable, i.e., $\Delta\omega_i \notin \theta$, and it is set to match the corresponding *symmetric* relative uncertainties — see (3).

Remark 3.1. If dealing with existing and functional positioning systems where prior knowledge does not exist or being inaccessible for some reason, then $k = 0$ in (10). Consequently, we start by having $N_T(t) = t$, and continue as described above.

3.3. Other applications

Interestingly, the proposed design framework of the trajectory signals can be useful in many other applications. For example, the suppress-excite signals can be utilized to control the mode shapes of a motion stage even without being over-actuated. Also, the designed suppress-excite signals with an appropriate choice of γ_f can be used for on-line fault detection, diagnosis, and machine health monitoring applications [66]. In this study, a single (S) input (I) and single (S) output (O), i.e., SISO, stage motion is handled. Dealing with other system input–output combinations, namely SIMO, MISO, and MIMO to obtain digital twins [67] of *multi-directional* motion stages based on data-driven modeling and controller synthesis [68], or first principles modeling [32] will be investigated in separate studies.

4. Simulation

Recalling Fig. 5 and based only on the available uncertain model of the motion stage (2), the Robust Control Toolbox™ [69] was used to robustly tune the tunable parameters $\{k_p, k_i, k_d, \tau_c, \tau_f\}$ when $n_f = 3$ and the plant nominal values were $\widetilde{m}_1 = 42.5$ kg, $\widetilde{m}_2 = 8$ kg, $\widetilde{k}_1 = 10$ N/m, $\widetilde{k}_2 = 7$ N/m, $\widetilde{c}_1 = 10$ N s/m, $\widetilde{c}_2 = 80$ N s/m, $\Delta m_{1,2} = 10\%$, $\Delta c_{1,2} = 50\%$, and $\Delta k_{1,2} = 40\%$. The resulting values of the parameters were $k_p = 468, k_i = 3.92 \times 10^5, k_d = 1.4 \times 10^5, \tau_c = 1 \times 10^{-4}$ and $t_f = 7.18 \times 10^{-5}$. Using 80 samples of the uncertain stage depicted in Fig. 5, the resonance frequencies associated with the lightly-damped modes in the feedforward and feedback paths are grouped into three frequency bands that are given as

$$
\begin{aligned}
\Omega_{s_1} &= \{\omega \in \mathbb{R} \mid 0.4445 - 0.1136 \le \omega \le 0.4445 + 0.1136\} \\
\Omega_{s_2} &= \{\omega \in \mathbb{R} \mid 1.6733 - 0.0127 \le \omega \le 1.6733 + 0.0127\} \\
\Omega_{s_3} &= \{\omega \in \mathbb{R} \mid 184.0130 - 59.0470 \le \omega \le 184.0130 + 59.0470\}
\end{aligned}
\tag{37}
$$

and should be suppressed to enhance the tracking performance. These frequency bands constitute Ω_s. Note that the feedforward filter introduces poles at high frequencies, i.e., 1.39×10^4 rad/s, and therefore its effect is ignored. Using (10) and (37), we find that $k = 3$, $\omega_{o_1} = 0.4445$, $\omega_{o_2} = 1.6733$, $\omega_{o_3} = 184.013$, and taking $\kappa_i = 2$ results in $\zeta_1^{min} = 0.1278$, $\zeta_2^{min} = 0.0038$, $\zeta_3^{min} = 0.1604$. Initially, we take $\zeta_i = \zeta_i^{min}, \forall i = 1, 2, 3$, which complies with Case 1.

Table 1
The coefficients of the excitation signal based on Method II rounded to 5 decimal places.

i	1	2	3	4
γ_{f_i}	0.00652	−0.01043	0.23932	1.52583
m_{s_i}	−2.13355	−7.11187	−795.35329	−1269.43723
θ_{s_i}	−0.21618	−0.7902	0.12202	−0.13609
D_i	1.20	0.30	0.10	0.10

Table 2
The coefficients of the simulated flexible modes according to (5), and their detected resonance frequencies.

i	$\tilde{\alpha}_i/m_t$	$\tilde{\zeta}_i$	$\tilde{\omega}_i$	Detected $\tilde{\omega}_i$
1	0.00019	0.05	49.00105	56.7465
2	$5e\times10^{-5}$	0.04973	82.88705	132
3	$5e\times10^{-5}$	0.00494	340.12382	350

Using (12) with $n=2$, we start building the suppress part of the unit-pulse acceleration signal according to (14), we then choose *an* optimal value of the pulse duration T through minimizing (27) with $\gamma_f=0$ and $\gamma_i=1, i=1,2,3$. According to Fig. 14(a), we choose $T=0.155$ s and ignore the other tiny value, i.e., $T\le 0.01$ s, because it corresponds to a high jerk value. Note that both Ω_{s_1} and Ω_{s_2} contribute most in (27) which may suggest choosing other values of the notch filters damping ratios. This can be verified in a later stage of the cyclic design process when the overall performance is evaluated.

Now, we locate Ω_s bands as depicted in Fig. 12, and we choose to limit the domain of interest to $\omega=140\,\pi$ rad/s. Consequently, the null-space of Ω_s within the interval $\omega\in[0,140\,\pi]$ is assigned to Ω_e, and therefore the excitation bands are specified according to (9) with $\Omega_{e_1}=0$. The results of designing the excitation part of the unit-pulse acceleration signal according to (18) using Method I and Method II are depicted in Figs. 14(b) through 14(d). Note that the excitation signal according to Method I fails to sweep Ω_{e_3}- see Fig. 14(c)-, and produced undesirable frequency contents in Ω_{e_2} where it ended up exciting Ω_{s_1}- see Fig. 14(b)- and violated the condition $\Omega_e\cap\Omega_s=\emptyset$. The results obtained using Method II are satisfactory and are used to realize the final unit-pulse suppress-excite acceleration signal with the coefficients values given in Table 1.

Before finalizing the suppress-excite signal design, the global excitation gain γ_f must be found. Recalling (26), we use $J_{max}=1600\,\mathrm{m/s^3}$ as the upper limit for the used jerk value in any generated profile with $U_{max}=15\,\mathrm{m/s^2}$ as an upper acceleration bound, as well. Based on (26), a maximum value of the *global* excitation gain, i.e., $\gamma_f=6750$, is found, which will ensure that the maximum jerk value generated by a unit-pulse acceleration signal will never exceed $106.667\approx(1600/15)\,\mathrm{m/s^3}$. Now, the unit-pulse suppress-excite signal can be designed according to (24) and Fig. 2.C. The resulting suppress-excite jerk, acceleration, and position profiles are depicted in Figs. 14(e) through 14(g), respectively, for $\gamma_f\in\{0,1000,6750\}$ that are used to schedule the transformation $\mathcal{T}$. Choosing $P_{T_1}=0.2$ m as the target point-to-point distance, both $\tilde{D}$ and a_1^{max} can be chosen appropriately such that $|a_1^{max}|\le 15\,\mathrm{m/s^2}$.

Using an existing digital twin [67,70] of the driven motion stage depicted in Fig. 5 that is uncertain with flexible modes, the resulting $\bar{\mathrm{M}}_2^2(\omega)$ measures for $\gamma_f\in\{0,1000,6750\}$ are depicted in Fig. 14(h), where 80-samples of the uncertain stage were averaged, and noise-free acceleration measurements of the output signal $p(t)$ are utilized. In the digital twin used, the flexible modes were according to (5) with $N=5$, and their coefficients are given in Table 2. The detected flexible modes frequencies given in Table 2 will be utilized to update mainly the suppress part of the input signal. Avoid exciting these modes will result in enhanced tracking as will be discussed shortly.

Let $\Omega_o=\{\omega_{o_1},\omega_{o_2},\omega_{o_3},19,25,38,57,132,350\}$ rad/s based on $\bar{\mathrm{M}}_2^2(\omega)$ measures depicted in Fig. 14(h). Using (36) to redefine mainly the suppress part of the acceleration signal, i.e., $\gamma_f=0$, with $\gamma_l=1,\forall l=1,2,\ldots,\bar{k}$, and $k=3$, $T=0.155$ s, and changing only ζ_3 as one example, and with $k=4$, and changing T, ω_{o_4}, ζ_i, $i=1,2,3,4$ as another example. This results in $\zeta_3=0.90$ for the first example, and in $T=0.15$ s, $\omega_{o_4}\approx 53.9114$ rad/s, $\zeta_1\approx 1.0435$, $\zeta_2\approx 0.4439$, $\zeta_3\approx 0.8007$, $\zeta_4\approx 0.500000007$ for the second example. The stage responses under these examples are depicted in Figs. 14(h) through 15(a), where the point-to-point (P2P) tracking performance improves significantly in the interval $t\in[0.4,0.75]$ s once compared to the one under the suppress acceleration signal without optimization in the same time interval. This resulted in two acceleration profiles that can maintain the desired tracking performance after introducing new modifications to the original stage without affecting the parameters of the involved control loops — see Remark 2.1. The overall picture of the whole framework that shows the interaction between the designed k-cascaded notch filter, control loops, and suppress/excite frequency bands is captured in Fig. 15(b), which shows that despite having a narrow band at $\omega=48.8$ rad/s in the closed-loop stage dynamics, the associated notch bandwidth, i.e., $\Delta\omega_4$, obtained through optimization is broad. This in fact reveals the nature of cascaded notch filters whose tabs individual properties are intermingled. Note also that the tabs frequencies, i.e., ω_{o_i}, have to be rearranged in ascending order to comply with (36) in case it should be used any further.

5. Conclusion

In this study, a framework that enables designing kinodynamical reference trajectories, and detecting the flexible modes of uncertain systems while in motion was proposed to enhance the tracking performance. According to this framework, there are two modes of operation, i.e., the suppress and suppress-excite modes. In the suppress mode of operation, the impulse response of

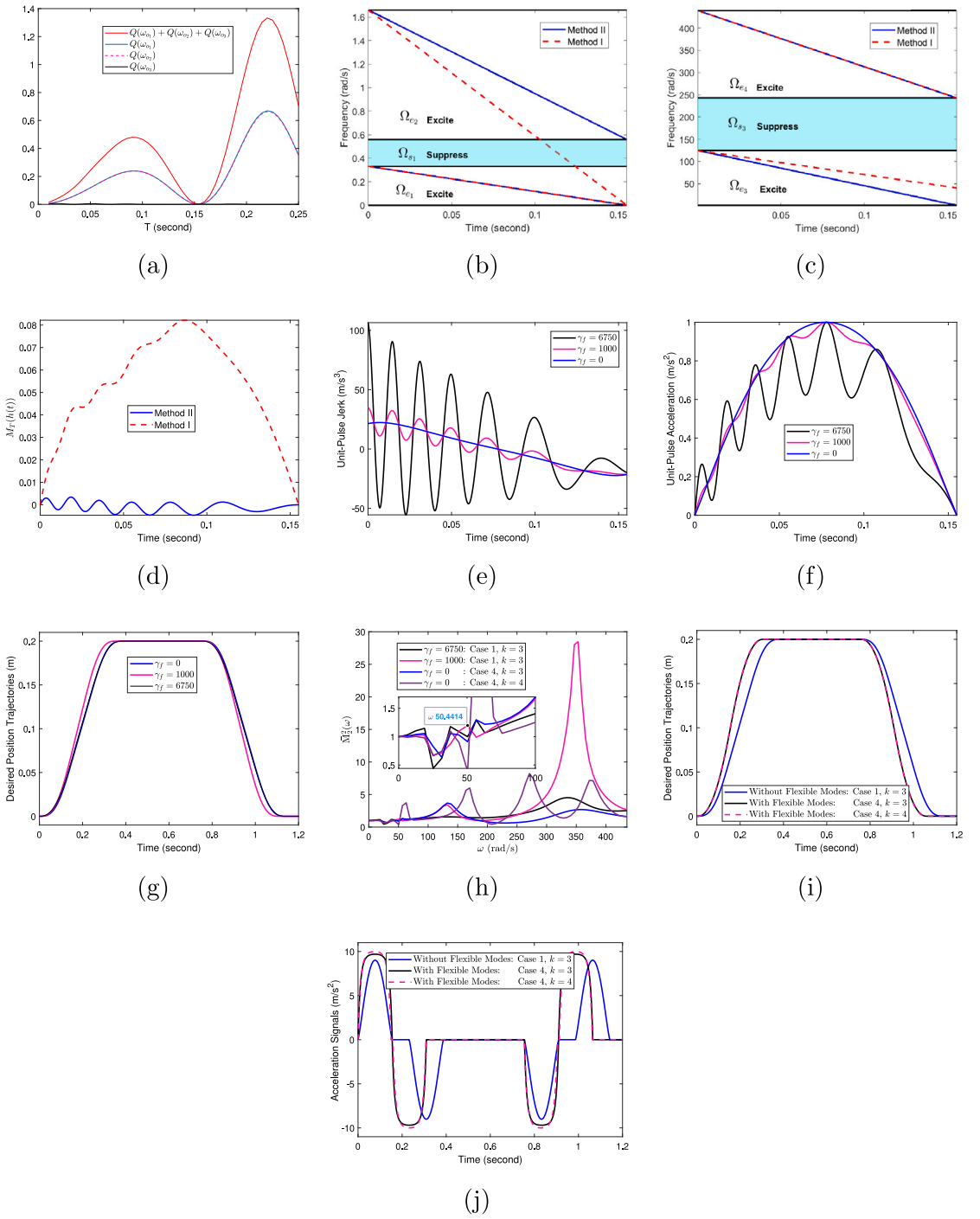

Fig. 14. (a) The magnitude squared of the *truncated* Fourier transform of the unit-pulse acceleration signal under Case 1. The generated excitation signals using the local and global design methods with their associated frequency bands (b) Ω_{e_1}, Ω_{e_2} and Ω_{s_1}, (c) Ω_{e_3}, Ω_{e_4} and Ω_{s_3}, and (d) the resulting $M_T(h(t))$ profiles. The tiny suppression band Ω_{s_2} is omitted. Based on the *modeled* part of the uncertain stage: the desired (e) unit-pulse jerk, (f) unit-pulse acceleration, (g) position trajectories, (h) The $\bar{M}_2^2$ variations when detecting the modal resonant frequencies of the flexible modes, and after updating the notch filter. Based on the *modeled* and *detected* parts of the uncertain stage: the desired (i) position trajectories, and (j) acceleration.

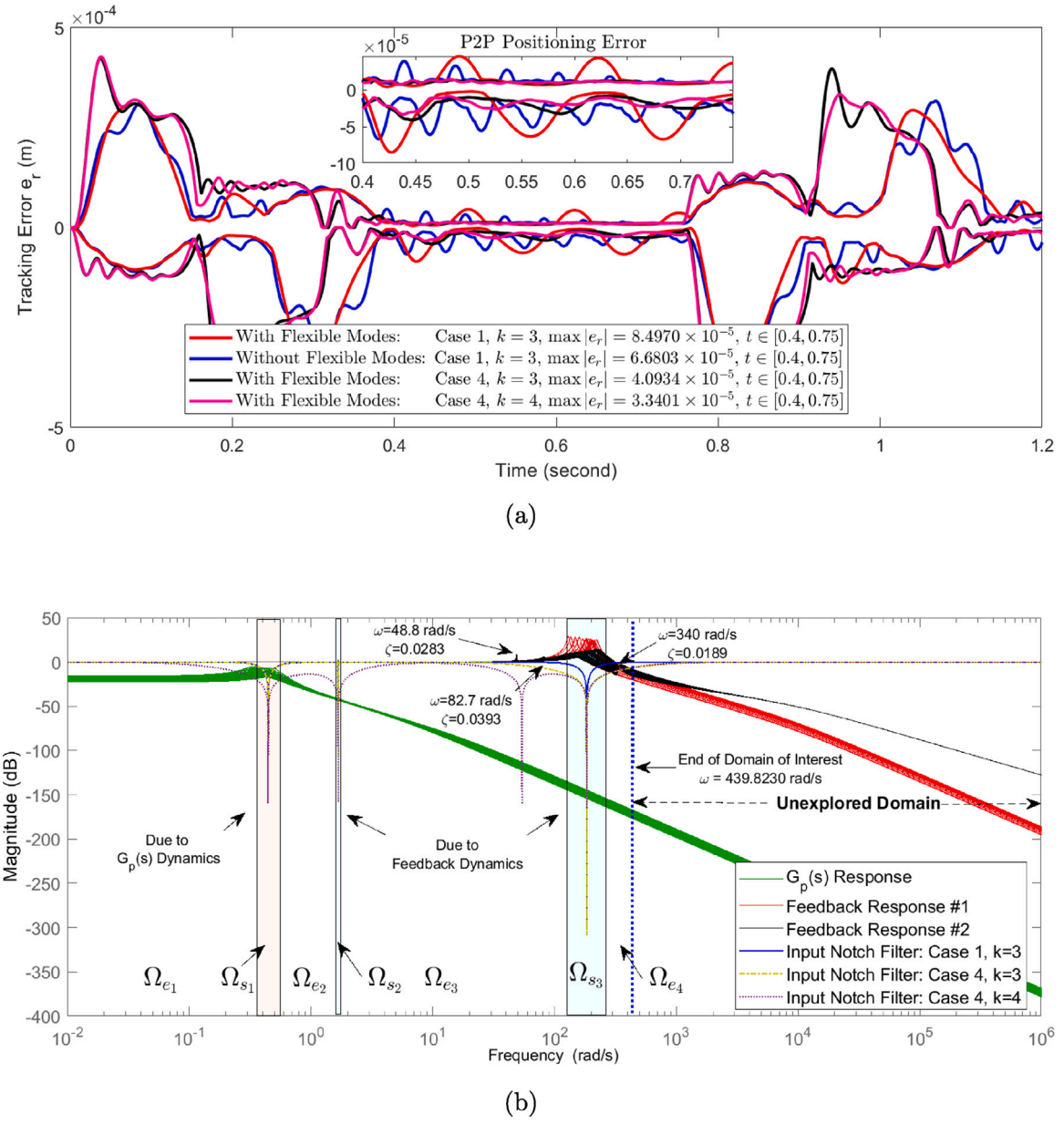

Fig. 15. (a) The tracking errors e_r envelops under the designed suppress signals. The blue lines indicate the envelops in the ideal case where no flexible modes are considered, while the other lines indicate the envelopes when the system including the flexible modes is simulated. The complete picture (b) of the suppress-excite signals given in the frequency domain where the interaction of the various components is made clear. The partially known uncertain stage dynamics $G_p(s)$ response is depicted in green. From $p_d(t)$ to $p(t)$, the feedback closed-loop *theoretical* response (#1) is depicted in red, and depends only on the *uncertain values* of the partially known stage dynamics, i.e., without including the flexible mode dynamics. Also, the feedback closed-loop *actual* response (#2) is depicted in black, and depends on both the *uncertain values and flexible modes* of the stage dynamics. Also, the three input notch filters used in this study are shown.

cascaded notch filters was used as a frequency-domain template to assign preferred frequency contents to a polynomial given by a time-domain template, and it is meant solely to suppress the known lightly-damped modes. In the suppress-excite mode of operation, excitation signals given in the form of constant amplitude linear chirp signals were superimposed on top of the suppress signal; to facilitate detecting possible lightly-damped modes in the frequency domain of interest. The utilization of the detected modes into the designed unit-pulse acceleration signal was achieved through optimization. Then, the acceleration amplitude needed to reach the target position was applied to the unit-pulse acceleration signal. According to the proposed framework, the deterioration of

the tracking performance due to the available flexible modes can be minimized using embedded input shaping – up to 39% in the provided example –, which does not require any modification to the controllers in the control loops.

Adopting a gray-box model of the positioning system, enables dealing with it from input–output perspective. Therefore, the proposed framework can operate on a functional positioning system with little knowledge about the system if any at all, which makes it suitable to be integrated within industrial or commercial motion controllers with auto-tuning features enabled. Basically, the suppression signal operates on a list of known frequencies associated with lightly-damped modes usually found in flexible uncertain systems, while the main role of the excitation signal is to expand that list by finding other lightly-damped modes. This also may result in increasing the operation bandwidth of the involved control loops, and increased machine throughput can be made possible.

Using function composition allowed using IIR filters in input shaping by creating a frequency-domain template with tunable parameters that can be adjusted to shape the time-domain template given as a polynomial of nth-order. Taking $n = 2$ allowed having analytical closed-form solution of the composed signal that is used to steer the motion stage. Moreover, trajectory extrapolation was made easy due to this analytical solution with reduced real-time calculation overhead. Since the function composition is a noncommutative process, reverse order of function composition will be the subject of future work.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

Acknowledgments

The authors acknowledge the financial support of the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[1] M. Heertjes, H. Butler, N. Dirkx, S. van der Meulen, R. Ahlawat, K. O'Brien, J. Simonelli, K. Teng, Y. Zhao, Control of wafer scanners: Methods and developments, in: 2020 American Control Conference, ACC, Denver,USA, 2020, pp. 3686–3703.
[2] H. Butler, Position control in lithographic equipment, IEEE Control Syst. Mag. 31 (5) (2011) 28–47.
[3] R. Schmidt, Ultra-precision engineering in lithographic exposure equipment for the semiconductor industry, Phil. Trans. R. Soc. A 370 (1973) (2012) 3950–3972.
[4] D. Amin-Shahidi, D. Trumper, Design and control of a piezoelectric driven reticle assist device for prevention of reticle slip in lithography systems, Mechatronics 24 (6) (2014) 562–571.
[5] B. Sencer, S. Tajima, Frequency optimal feed motion planning in computer numerical controlled machine tools for vibration avoidance, J. Manuf. Sci. Eng. 139 (1) (2017).
[6] Y. Altintas, Manufacturing Automation: Metal Cutting Mechanics, Machine Tool Vibrations, and CNC Design, second ed., Cambridge University Press, 2012.
[7] D. Bruijnen, N. van Dijk, Combined input shaping and feedforward control for flexible motion systems, in: 2012 American Control Conference, ACC, Montreal, Canada, 2012, pp. 2473–2478.
[8] F. Boeren, D. Bruijnen, N. van Dijk, T. Oomen, Joint input shaping and feedforward for point-to-point motion: Automated tuning for an industrial nanopositioning system, Mechatronics 24 (6) (2014) 572–581, Control of High-Precision Motion Systems.
[9] N. Singer, W. Seering, Preshaping command inputs to reduce system vibration, 1990.
[10] B. Sencer, K. Ishizaki, E. Shamoto, High speed cornering strategy with confined contour error and vibration suppression for CNC machine tools, CIRP Ann. 64 (1) (2015) 369–372.
[11] T. Oomen, R. van Herpen, S. Quist, M. van de Wal, O. Bosgra, M. Steinbuch, Connecting system identification and robust control for next-generation motion control of a wafer stage, IEEE Trans. Control Syst. Technol. 22 (1) (2014) 102–118.
[12] A. Fleming, A. Wills, Optimal periodic trajectories for band-limited systems, IEEE Trans. Control Syst. Technol. 17 (3) (2009) 552–562.
[13] L. Piegl, W. Tiller, The NURBS Book, Springer Science & Business Media, Berlin/Heidelberg, Germany, 1996.
[14] B. Sencer, Y. Kakinuma, Y. Yamada, Linear Interpolation of machining tool-paths with robust vibration avoidance and contouring error control, Precis. Eng. 66 (2020) 269–281.
[15] K. Erwinski, A. Wawrzak, M. Paprocki, Real-time jerk limited feedrate profiling and interpolation for linear motor multi-axis machines using NURBS toolpaths, IEEE Trans. Ind. Inf. 18 (2022) 7560–7571.
[16] M. Lin, M. Tsai, H. Yau, Development of a dynamics-based NURBS interpolator with real-time look-ahead algorithm, Int. J. Mach. Tools Manuf. 47 (15) (2007) 2246–2262.
[17] T. Vyhlídal, M. Hromčík, Parameterization of input shapers with delays of various distribution, Automatica 59 (2015) 256–263.
[18] Y. Al-Rawashdeh, M. Al Janaideh, M. Heertjes, On step-and-scan trajectories used in wafer scanners in semiconductor manufacturing, in: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems, Prague, Czech Republic, pp. 1–7.
[19] Y. Fang, J. Hu, W. Liu, Q. Shao, J. Qi, Y. Peng, Smooth and time-optimal S-curve trajectory planning for automated robots and machines, Mech. Mach. Theory 137 (2019) 127–153.
[20] P. Lambrechts, M. Boerlage, M. Steinbuch, Trajectory planning and feedforward design for electromechanical motion systems, Control Eng. Pract. 13 (2) (2005) 145–157.
[21] K. Chen, R. Fung, The point-to-point multi-region energy-saving trajectory planning for a mechatronic elevator system, Appl. Math. Model. 40 (21–22) (2016) 9269–9285.
[22] J. Lawrence, W. Singhose, Command shaping slewing motions for tower cranes, J. Vib. Acoust. 132 (1) (2010).
[23] C. Liu, M. Tsai, M. Hong, P. Tang, Development of a novel tuning approach of the notch filter of the servo feed drive system, J. Manuf. Mater. Process. 4 (1) (2020) 21.

[24] W. Singhose, R. Eloundou, J. Lawrence, Command generation for flexible systems by input shaping and command smoothing, J. Guid. Control Dyn. 33 (6) (2010) 1697–1707.
[25] W. Singhose, S. Derezinski, N. Singer, Input shapers for improving the throughput of torque-limited systems, in: Proceeding of the 1994 Conference on Control Applications, 1994.
[26] G. Song, N. Buck, B. Agrawal, Spacecraft vibration reduction using pulse-width pulse-frequency modulated input shaper, J. Guid. Control Dyn. 22 (3) (1999) 433–440.
[27] G. Lee, J. Kim, Y. Choi, Convolution-based trajectory generation methods using physical system limits, J. Dyn. Syst. Meas. Control 135 (1) (2013) 011001.
[28] B. Donald, P. Xavier, J. Canny, J. Reif, Kinodynamic motion planning, J. ACM 40 (5) (1993) 1048–1066.
[29] W. Obi, Error analysis of a Laplace transform inversion procedure, SIAM J. Numer. Anal. 27 (2) (1990) 457–469.
[30] K. Kuhlman, Review of inverse Laplace transform algorithms for Laplace-space numerical approaches, Numer. Algorithms 63 (2) (2013) 339–355.
[31] S. Bergner, T. Moller, D. Weiskopf, D. Muraki, A spectral analysis of function composition and its implications for sampling in direct volume visualization, IEEE Trans. Vis. Comput. Graphics 12 (5) (2006) 1353–1360.
[32] Y. Al-Rawashdeh, M. Al Janaideh, M. Heertjes, On characterization of a generic lithography machine in a multi-directional space, Mech. Mach. Theory 170 (2022) 104638.
[33] L. Dai, X. Li, Y. Zhu, M. Zhang, C. Hu, The generation mechanism of tracking error during acceleration or deceleration phase in ultraprecision motion systems, IEEE Trans. Ind. Electron. 66 (9) (2018) 7109–7119.
[34] K. Verkerk, H. Butler, P. van den Bosch, Improved disturbance rejection for high precision systems through estimation of the flexible modes, in: 2015 IEEE Conference on Control Applications, CCA, Sydney, Australia, 2015, pp. 1191–1196.
[35] M. Ronde, M. Schneiders, E. Kikken, M. Van De Molengraft, M. Steinbuch, Model-based spatial feedforward for over-actuated motion systems, Mechatronics 24 (4) (2014) 307–317.
[36] R. Subramanian, M. Heertjes, T. de Hoog, A model-based inferential feedforward approach to deal with hysteresis in a motion system, in: 2018 Annual American Control Conference, ACC, Milwaukee, USA, 2018, pp. 2934–2939.
[37] M. Hoogerkamp, R. Waiboer, R. Aarts, Modeling of flexible non-linear dynamic links in nano-positioning motion systems, in: 9th ECCOMAS Thematic Conference on Multibody Dynamics, ECCOMAS, pp. 1–14.
[38] B. Xia, C. Yuan, Y. Tian, S. Wu, K. Yang, Disturbance estimation and compensation for planar motors on the long-stroke stage of a wafer stage, Adv. Mech. Eng. 7 (4) (2015) 1–10.
[39] S. Devasia, Should model-based inverse inputs be used as feedforward under plant uncertainty? IEEE Trans. Automat. Control 47 (11) (2002) 1865–1871.
[40] Y. Kasemsinsup, M. Heertjes, H. Butler, S. Weiland, Exact plant inversion of flexible motion systems with a time-varying state-to-output map, in: 2016 European Control Conference, ECC, Aalborg, Denmark, 2016, pp. 2483–2488.
[41] B. Rigney, L. Pao, D. Lawrence, Settle time performance comparisons of stable approximate model inversion techniques, in: 2006 American Control Conference, Minneapolis, USA, 2006, pp. 600–605.
[42] K. Leang, Q. Zou, S. Devasia, Feedforward control of piezoactuators in atomic force microscope systems, IEEE Control Syst. Mag. 29 (1) (2009) 70–82.
[43] G. Clayton, S. Tien, K. Leang, Q. Zou, S. Devasia, A review of feedforward control approaches in nanopositioning for high-speed SPM, J. Dyn. Syst. Meas. Control 131 (6) (2009) 061101.
[44] G. Ferretti, G. Magnani, P. Rocco, Some fundamental limitations in the control of two-mass systems, in: 2009 IEEE International Conference on Mechatronics, Malaga, Spain, 2009, pp. 1–6.
[45] B. Fine, S. Mishra, M. Tomizuka, Model inverse based iterative learning control using finite impulse response approximations, in: 2009 American Control Conference, St. Louis, USA, 2009, pp. 931–936.
[46] V. Nozhenko, D. Rodkin, K. Bohatyrov, Control of passing the resonance zone during start-up of above resonance vibration machines, in: 2019 IEEE International Conference on Modern Electrical and Energy Systems, MEES, Kremenchuk, Ukraine, 2019, pp. 146–149.
[47] K. Geldhof, A. Van den Bossche, J. Melkebeek, Influence of electrical eigenfrequencies on damped voltage resonance based sensorless control of switched reluctance drives, in: 2009 35th Annual Conference of IEEE Industrial Electronics, Porto, Portugal, 2009, pp. 4058–4064.
[48] G. Ellis, R. Lorenz, Resonant load control methods for industrial servo drives, in: Conference Record of the 2000 IEEE Industry Applications Conference. Thirty-Fifth IAS Annual Meeting and World Conference on Industrial Applications of Electrical Energy (Cat. No. 00CH37129), Vol. 3, Rome, Italy, 2000, pp. 1438–1445.
[49] T. Miyoshi, Y. Noda, K. Terashima, Feedforward control considering input and states constraints with eliminating residual vibration, in: 2007 American Control Conference, New York, USA, 2007, pp. 5005–5010.
[50] Residue MATLAB function for partial fraction expansion, 2022, https://www.mathworks.com/help/matlab/ref/residue.html. Accessed: 2022-05-14.
[51] Q. Wang, D. Kundur, A generalized design framework for IIR digital multiple notch filters, EURASIP J. Adv. Signal Process. 2015 (1) (2015) 1–13.
[52] Y. Fang, J. Hu, W. Liu, Q. Shao, J. Qi, Y. Peng, Smooth and time-optimal S-curve trajectory planning for automated robots and machines, Mech. Mach. Theory 137 (2019) 127–153.
[53] D. Eager, A. Pendrill, N. Reistad, Beyond velocity and acceleration: jerk, snap and higher derivatives, Eur. J. Phys. 37 (6) (2016) 065008.
[54] G. Milton, J. Willis, On modifications of Newton's second law and linear continuum elastodynamics, Proc. R. Soc. A 463 (2079) (2007) 855–880.
[55] J. Banerjee, Frequency dependent mass and stiffness matrices of bar and beam elements and their equivalency with the dynamic stiffness matrix, Comput. Struct. 254 (2021) 106616.
[56] H. Vold, J. Crowley, G. Rocklin, New ways of estimating frequency response functions, Sound Vib. 18 (11) (1984) 34–38.
[57] A. Kallel, D. Bouchaala, O. Kanoun, Critical implementation issues of excitation signals for embedded wearable bioimpedance spectroscopy systems with limited resources, Meas. Sci. Technol. 32 (8) (2021) 084011.
[58] F. Saupe, A. Knoblach, Experimental determination of frequency response function estimates for flexible joint industrial manipulators with serial kinematics, Mech. Syst. Signal Process. 52 (2015) 60–72.
[59] J. Vuojolainen, N. Nevaranta, R. Jastrzebski, O. Pyrhönen, Comparison of excitation signals in active magnetic bearing system identification, Model. Identif. Control 38 (3) (2017) 123–133.
[60] X. Shan, H. Song, C. Zhang, G. Wang, J. Fan, Linear system identification and vibration control of end-effector for industrial robots, Appl. Sci. 10 (23) (2020) 8537.
[61] A. Novak, L. Simon, F. Kadlec, P. Lotton, Nonlinear system identification using exponential swept-sine signal, IEEE Trans. Instrum. Meas. 59 (8) (2009) 2220–2229.
[62] X. Xia, System identification using chirp signals and time-variant filters in the joint time-frequency domain, IEEE Trans. Signal Process. 45 (8) (1997) 2072–2084.
[63] R. Lederman, V. Rokhlin, On the analytical and numerical properties of the truncated Laplace transform I, SIAM J. Numer. Anal. 53 (3) (2015) 1214–1235.
[64] M. Sandoval-Hernandez, H. Vazquez-Leal, L. Hernandez-Martinez, U. Filobello-Nino, V. Jimenez-Fernandez, A. Herrera-May, R. Castaneda-Sheissa, R. Ambrosio-Lazaro, G. Diaz-Arango, Approximation of Fresnel integrals with applications to diffraction problems, Math. Probl. Eng. 2018 (2018).
[65] H. Jeffreys, Weierstrass's theorem on approximation by polynomials" and" Extension of Weierstrass's approximation theory, in: Methods of Mathematical Physics, Cambridge University Press, 1988, pp. 446–448.
[66] R. Voorhoeve, R. de Rozario, W. Aangenent, T. Oomen, Identifying position-dependent mechanical systems: A modal approach applied to a flexible wafer stage, IEEE Trans. Control Syst. Technol. 29 (1) (2021) 194–206.

[67] K. Classens, W. Heemels, T. Oomen, Digital twins in mechatronics: From model-based control to predictive maintenance, in: 2021 IEEE 1st International Conference on Digital Twins and Parallel Intelligence, DTPI, Beijing, China, 2021, pp. 336–339.
[68] T. Bloemers, T. Oomen, R. Tóth, Frequency response data-driven LPV controller synthesis for MIMO systems, IEEE Control Syst. Lett. (2021) 1.
[69] Robust Control Toolbox description, 2021, https://www.mathworks.com/products/robust.html. Accessed: 2021-07-05.
[70] E. Negri, L. Fumagalli, M. Macchi, A review of the roles of digital twin in CPS-based production systems, in: Procedia Manufacturing, 11, 2017, pp. 939–948.