Robust broadband nanopositioning: fundamental trade-offs, analysis, and design in a two-degree-of-freedom control framework

2009 Nanotechnology 20 035501

(http://iopscience.iop.org/0957-4484/20/3/035501)

Nanotechnology **20** (2009) 035501 (16pp) doi:10.1088/0957-4484/20/3/035501

Robust broadband nanopositioning: fundamental trade-offs, analysis, and design in a two-degree-of-freedom control framework

Chibum Lee and Srinivasa M Salapaka

Department of Mechanical Sciences and Engineering, University of Illinois at Urbana-Champaign, 1206 West Green Street, Urbana, IL 61801-2906, USA

E-mail: salapaka@illinois.edu

Received 31 March 2008, in final form 9 November 2008
Published 16 December 2008
Online at stacks.iop.org/Nano/20/035501

Abstract

This paper studies and analyses fundamental trade-offs between positioning resolution, tracking bandwidth, and robustness to modeling uncertainties in two-degree-of-freedom (2DOF) control designs for nanopositioning systems. The analysis of these systems is done in optimal control setting with various architectural constraints imposed on the 2DOF framework. In terms of these trade-offs, our analysis shows that the primary role of feedback is providing robustness to the closed-loop device whereas the feedforward component is mainly effective in overcoming fundamental algebraic constraints that limit the feedback-only designs. This paper presents (1) an optimal prefilter model matching design for a system with an existing feedback controller, (2) a simultaneous feedforward and feedback control design in an optimal mixed sensitivity framework, and (3) a 2DOF optimal robust model matching design. The experimental results on applying these controllers show a significant improvement, as high as 330% increase in bandwidth for similar robustness and resolution over optimal feedback-only designs. Other performance objectives can be improved similarly. We demonstrate that the 2DOF freedom design achieves performance specifications that are analytically *impossible* for feedback-only designs.

(Some figures in this article are in colour only in the electronic version)

1. Introduction

The last two decades have seen a rapid development in nanoscientific research and related nanotechnology. Recent advances in interrogation, control, and manipulation of various properties of matter at the atomic scale, evidenced by devices such as atomic force microscopes (AFMs) and related scanning-probe microscopes (SPMs), are having a dramatic impact on fields as diverse as biology, materials science, electrochemistry, tribology, biochemistry, surface physics, and medicine [1]. Apart from their impact on science, these advances are fueling new technologies. For example, the ability to manipulate and sense the topography of material at the nanoscale has resulted in new technologies for high-density data storage [2, 3]. Similarly, the ability to modify material biologically is leading to bio-molecular assays that can be used for drug discovery [4]. However, many hurdles still have to be overcome before the goal of *practical, reliable, and routine* control and manipulation of matter at the atomic scale is realized.

One of the pivotal requirements of nanotechnology is nanopositioning. Nanoscientific studies as well as applications, such as in scanning-probe microscopy, demand positioning systems with atomic scale resolutions. There is an added impetus on the design of nanopositioning systems since they form the bottleneck in terms of speed and accuracy of most devices for nano-investigation, especially in SPMs. For instance in AFMs, the positioning resolution and tracking bandwidth of positioning systems is typically a few orders less than the imaging resolution and bandwidth that a

microcantilever probe provides. Besides high-precision positioning, most nanoscientific studies and applications impose severe demands on the tracking bandwidth and reliability in terms of repeatability of experiments. High tracking bandwidth is required as many studies, especially in biology and material science, require assaying matter with nanoscale precision over areas with characteristic lengths that are typically three orders or more. Repeatability of experiments is essential for validation of the underlying studies. This requirement translates to robustness of positioning systems to modeling uncertainties and operating conditions. Devices that are insensitive to (robust to) diverse operating conditions give repeatable measurements, and are hence reliable.

Typical nanopositioning systems comprise a flexure stage that provides frictionless motion through elastic deformation, and an actuator, typically made from piezoelectric material that provides the required force to deform the flexure stage and/or sensing system along with the control system. The main challenges to design of robust broadband nanopositioning systems come from flexure-stage dynamics that limit the bandwidth of the positioning stage, from nonlinear effects of piezoactuation such as hysteresis and creep that are difficult to model, and from sensor noise management issues in control feedback that can potentially hamper the tracking resolution of the device.

The most common approach to increasing tracking bandwidth is to increase the resonance frequency of a nanopositioning system by building devices that are sufficiently stiff and small. However, building small scanning devices limits the maximum traversal of the system to a few microns. A feedforward scheme for improving the accuracy of a nanopositioning system is discussed in [5]. Energy-based models of piezoelectric behavior that incorporate the dependence of hysteresis loops on the scan rates are developed and applied to piezoelectric actuators in [6, 7]. These models, which are based on physical principles, predict a greater number of observed features in experiments than the phenomenological models, such as the Preisach model [8, 5]. Feedforward implementation schemes based on these models thus provide better positioning resolution [6]. The use of charge amplifiers instead of voltage amplifiers is another way for reducing hysteresis [9].

Feedback control designs [10, 11] with large gains at low frequencies have been demonstrated that make positioning resolution practically independent of piezoelectric nonlinearities, where nonlinear effects become negligible compared to measurement noise. Furthermore, appropriate feedback designs are less sensitive to operating conditions than their feedforward counterparts. The feedback control framework presented in [11] determines and quantifies trade-offs between performance objectives, assesses if desired specifications are feasible, and provides a way to design controllers to achieve specifications when possible. More recently, two-degree-of-freedom (2DOF) designs that combine the feedforward and feedback strategies have been reported [12–15]. In [13], a 2DOF design is presented where the regular feedback control is appended with a feedforward scheme that exploits the information from the previous scan line to improve the scanning performance for imaging the current scan line. In [14], 2DOF design schemes are discussed where optimal inversion control (as well as iterative learning schemes [16]) are used along with the feedback control to get improvements on the feedback-only design. In [15], a polynomial-based feedback controller is designed to account for pole-location uncertainty along with an inversion-based feedforward controller which provides improved tracking for raster-scan applications.

This paper presents 2DOF control design for nanopositioning systems in an optimal control framework. The main contributions of this paper are threefold. First, it characterizes fundamental limitations on the space of achievable performance specifications in the 2DOF scheme. These limitations are fundamental in the sense that they are completely determined by the device design and are independent of the control design. Based on these limitations, this paper quantifies and analyses the trade-offs between positioning resolution, tracking bandwidth, and robustness to modeling uncertainties in the 2DOF design scheme. This paper discusses the extent of increase in the space of feasible performance specifications in the 2DOF design scheme when compared to feedback-only designs. Second, it poses optimization problems for determining *simultaneously* both the feedback and feedforward control laws to achieve design specifications that are within the fundamental constraints. It presents the optimal 2DOF control design for three practical scenarios — when the feedback design is given *a priori*, when performance responsibilities between the feedback and feedforward controls are decided *a priori*, and when neither feedback design is given nor the responsibility-sharing given *a priori*. Third, it presents an analysis and comparison of the roles of the feedforward and feedback components of 2DOF design, and studies the extent of improvements over the feedback-only designs. Theoretical assertions are substantiated by experimental data. In particular, experimental results are provided that prove that 2DOF designs do achieve performance specifications which no feedback-only design can achieve.

The paper is organized as follows. Performance characterization of nanopositioning systems and the objectives of and limitations on their control are given in section 2. A 2DOF control framework is introduced and three control design methods are presented in section 3. In section 4, these 2DOF control designs are demonstrated on an AFM scanner and substantiated through experimental results. The analysis and discussion of this framework is presented in section 5. Some observations and conclusions are presented in section 6.

2. Performance characterization, objectives, and limitations

A typical nanopositioning system used in an SPM is comprised of a flexure stage, and actuators (typically piezoelectric) and/or sensors along with the feedback system. We present our analysis and design in terms of transfer function block diagrams as shown in figure 1. The transfer function of a system represents the linear dynamics of the system about an operating point. In this figure, for instance, G is the transfer

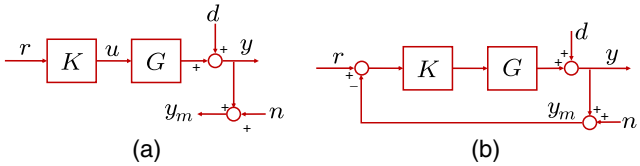

Figure 1. Block diagram schematics for nanopositioning systems. The controller acts only on the reference input signal in open-loop nanopositioning systems shown in (a) while it has access to the difference between the reference and the position y in the closed-loop (feedback-only) positioning system as shown in (b).

function of the *scanner* comprising the actuator, the flexure stage, and the sensor. It represents the dynamical relationship between its output, the flexure-stage displacement y (which is scaled by a sensor constant), and its input, the voltage u given to the actuator. The relation is denoted by a product operation as $y = Gu$ (or as $y = G(\mathrm{j}\omega)u$ when the relation between the Fourier components of frequency ω is emphasized). In figure 1, the signal r represents the command signal that the positioning system needs to track, the disturbance signal d represents the *mechanical noise*—the effects of dynamics that are not incorporated in the model G, n represents the sensor noise, $y_{\mathrm{m}} = y+n$ represents the noisy measurement signal, and the transfer function K represents the control transfer function. The main objective for the design of the controller K is to make the *tracking error* small, that is to make the difference $r - y$ between the desired and actual motions small.

In open-loop positioning systems (figure 1(a)), where the sensor signal is not fed back to the controller, the performance is severely limited by mechanical noise. The mechanical noise mainly consists of the slowly varying drift and creep, which are therefore prominent in slow scans, and the inertial lag at high frequencies, which is prominent in high-speed scans. Hysteresis affects the systems at all frequencies and is particularly prominent in repetitive raster scanning. These nonlinear effects of drift, creep, and hysteresis are sensitive to changes in operating conditions such as ambient temperature, residual polarization in piezoactuators, and most importantly the operating point, the reference value on the nonlinear input–output (input voltage versus stage displacement) graph about which stage motions are calibrated. Therefore, including their precise behavior in device models is practically infeasible, and hence they are treated as noise. Feedback-based schemes (figure 1(b)) have demonstrated effective compensation for the creep, drift, hysteresis, and inertial lag problems without requiring their precise models [11]. They compensate for the mechanical noise but at the cost of feeding back the relatively smaller sensor noise. In this paper, we do not consider open-loop systems and analyze systems that have feedback along with feedforward components. An advantage of block diagram schematics as in figure 1(b) is that objectives and limitations on the positioning systems can be quantified in terms of the transfer functions for closed-loop systems. We first present objectives and limitations in feedback-only designs and then present the 2DOF framework.

2.1. Performance characterization and control objectives

The performance of a nanopositioning system is characterized by its positioning resolution, tracking bandwidth, and robustness to modeling uncertainties. The resolution of the nanopositioning system is specified in terms of the standard deviation σ of the sensor output when there is no actuation of the positioning stage. The measurement noise typically exhibits a zero-mean Gaussian distribution. Thus, 3σ-resolution defined by the measurement noise gives over 99.7% confidence in any signal value that is greater than the resolution. The tracking bandwidth is the range of reference-signal frequencies that the nanopositioning system can track with a given precision. To characterize the robustness to modeling uncertainties, a metric is needed that quantifies how insensitive the closed-loop device is to the modeling errors and operating conditions.

In feedback-only configuration (figure 1(b)), these performance specifications can be quantified by analyzing the tracking error. For a given controller K, the tracking error in this configuration is given by

$$e = r - y = S(r - d) + Tn, \tag{1}$$

where the *sensitivity transfer function* $S = (1 + GK)^{-1}$ and the *complementary sensitivity transfer function* $T = 1 - S = (1+GK)^{-1}GK$. Thus, the tracking error, which is given by the superposition of the error contribution Sr from the reference signal r and the error contribution Tn from the sensor noise n, is adversely affected by large values of scan range of r and the amplitude of noise n. Thus, high resolution can be achieved by designing the feedback law $K(\mathrm{j}\omega)$ such that $S(\mathrm{j}\omega)$ and $T(\mathrm{j}\omega)$ are small in those frequency ranges where the frequency contents of r and n, respectively, are large. The resolution of the closed-loop positioning system is determined by the term Tn and therefore lower values of $T(\mathrm{j}\omega)$ over larger ranges of frequencies guarantee better resolutions. More specifically, the standard deviation σ of the zero-mean position signal when the reference signal is identically zero, which determines the resolution of the positioning system, is given by

$$\sigma = \int_0^\infty |T(\mathrm{j}\omega)|^2 P_n(\omega)\,\mathrm{d}\omega, \tag{2}$$

where $P_n(\omega)$ denotes the power spectral density of the noise signal n. Thus the smaller the bandwidth of T, which is characterized by the roll-off frequency ω_T (figure 2(a)), the smaller the standard deviation σ, and hence the better the resolution of the closed-loop device.

Similarly, in this configuration, the tracking bandwidth is determined by the range of frequencies over which the magnitude plot of the sensitivity transfer function $S(\mathrm{j}\omega)$ is below $1/\sqrt{2}$. It is characterized by the corner frequency ω_{BW} as shown in figure 2(b). Note that, unlike typical usage, we use ω_{BW} and not ω_T as a measure of the bandwidth. Even though ω_T gives a higher numerical value for the bandwidth for the same device, it is poorly related to the actual time constants of the device [17].

The sensitivity function $S(\mathrm{j}\omega)$ is so called since it is equal to $\frac{\mathrm{d}y/y}{\mathrm{d}G/G}$, that is it gives the percentage change in the

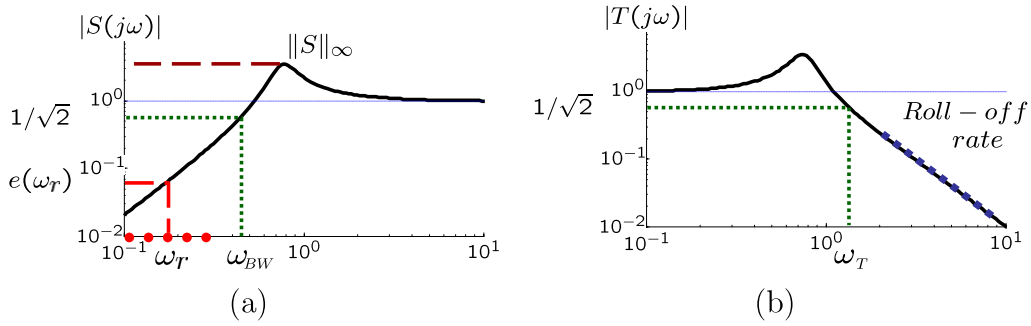

Figure 2. Objectives of feedback control. The objective of the control design is to achieve high positioning resolution (by achieving low values of ω_T, high roll-off rates in (a) and small error at reference frequencies $e(\omega_r)$ in (b)), high tracking bandwidth (by designing for large ω_{BW}) and robustness (by achieving $\|S\|_\infty$ close to 1).

position signal for a given percentage change in the model; therefore it serves as a measure of 'sensitivity' or robustness of the closed-loop loop device to modeling uncertainties. We use the peak value of the magnitude of sensitivity function, $\|S\|_\infty$, to characterize the robustness of the system to modeling uncertainties and operating conditions (figure 2(b)). Thus the performance specifications translate to control design objectives of achieving high values of ω_{BW} for high tracking bandwidth, high roll-off rates of T, and smaller values of ω_T for better positioning resolution, and low values of $\|S\|_\infty$ for better robustness to modeling uncertainties.

2.2. Limitations

These objectives have to be achieved under some practical and fundamental algebraic limitations. For instance, it is impractical to have the sampling period faster than the time required by the digital signal processor (DSP) to calculate the steps in the control logic. Thus both high-frequency sampling and high-order control logic cannot be implemented simultaneously. Similarly, to avoid errors due to the digital implementation of control laws, sampling frequencies around 30 times the frequency range of interest are typically prescribed [18]. Another important practical limitation is that the actuation signal should be within saturation limits of the hardware.

The algebraic limitations on the control design of the positioning systems are *fundamental* since they are completely determined by the scanner G and are independent of the controller K. For instance, in feedback-only systems $S(\mathrm{j}\omega) + T(\mathrm{j}\omega) \equiv 1$, which is evident from the definitions of S and T. This algebraic limitation prevents the error $e = r - y = Sr - Sd + Tn$ from becoming small in all frequencies since S and T cannot be made small simultaneously. This motivates the search for control designs that achieve a trade-off between the bandwidth and resolution requirements. Besides, for a scanner G with phase margin less than 90°, which is true for most practical systems, the bandwidth ω_{BW} cannot be larger than ω_T [17]. This limitation prevents the feedback control from achieving noise attenuation over the target reference frequency range. Another fundamental limitation that imposes a trade-off between the bandwidth, the resolution, and the robustness requirements can be explained in terms of the Bode integral law [19, 20]. This law states that for any stable system G such that the relative degree of the transfer function $K(s)G(s)$ is at least two,

$$\int_0^\infty \log|S(\mathrm{j}\omega)|\,\mathrm{d}\omega = 0. \tag{3}$$

Remark. The condition on the relative degree, which is the difference of the orders of the denominator and numerator polynomials in the transfer function [17], is typically satisfied. More specifically, the open-loop transfer function $K(s)G(s)$ is designed such that it has relative degree of order greater than or equal to 2 in order to ensure the closed-loop transfer function T has a sufficiently fast roll-off rate at high frequencies for noise attenuation (and therefore better resolution). Also, when a discrete control system is used, this relative degree condition is inherently satisfied [21]. □

The limitations from this algebraic law can be explained in terms of the *waterbed effect*—since the area under the graph of $\log|S(\mathrm{j}\omega)|$ over the entire frequency range is zero, $S(\mathrm{j}\omega)$ made small at a certain frequency range has to be compensated by making it large at some other frequency ranges. One direct consequence of this law is that $S(\mathrm{j}\omega)$ cannot be made less than 1 over all frequencies. Therefore $\|S\|_\infty$, the measure for robustness, is at least 1. Moreover, if the stable system G has real non-minimum phase zeros (that is real and positive roots), a stricter fundamental algebraic law holds:

$$\int_0^\infty \log|S(\mathrm{j}\omega)|\; W(z,\omega)\,\mathrm{d}\omega = 0, \tag{4}$$

where $W(z,\omega) = \frac{2z}{z^2+\omega^2}$ for real positive zero z [20].

Remark. Typical scanner systems have non-collocated actuators and sensors that are separated by flexure stages. The transfer function models of such systems generally exhibit non-minimum phase zeros. □

In this case it can be shown that the integral of $\log|S(\mathrm{j}\omega)|$ over a *finite* frequency region can be bounded from below, thus manifesting a waterbed effect over a finite frequency range. Thus the simultaneous requirements of low $|S(\mathrm{j}\omega)|$ over a large frequency for a high tracking bandwidth, high-order roll-off rates of T at high frequencies for high resolution, and small peaks of the $S(\mathrm{j}\omega)$ compete against each other under

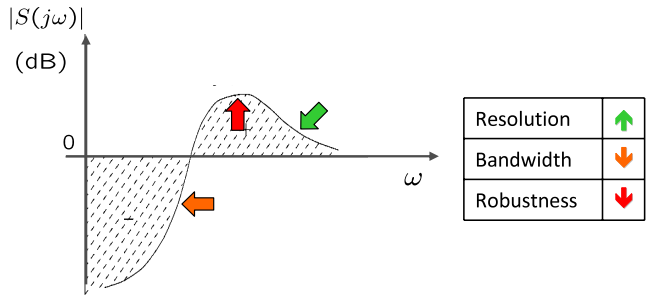

Figure 3. Trade-offs due to the finite-waterbed effect. The Bode integral laws manifest themselves as waterbed effects, where decreasing the magnitude of sensitivity function at a certain frequency range results in its increase in some other frequency range. For instance, making the sensitivity function small (near to 1) at high frequencies to ensure a high roll-off rate of the complementary sensitivity function (since $T = 1 - S$) for better resolution results in lower robustness to modeling uncertainties (due to higher values of the peak ($\|S\|_\infty$)) and lower values of tracking bandwidth (since ω_{BW} decreases). Similar trade-offs where one performance objective is sacrificed at the cost of others can be analyzed by studying the finite-waterbed effect.

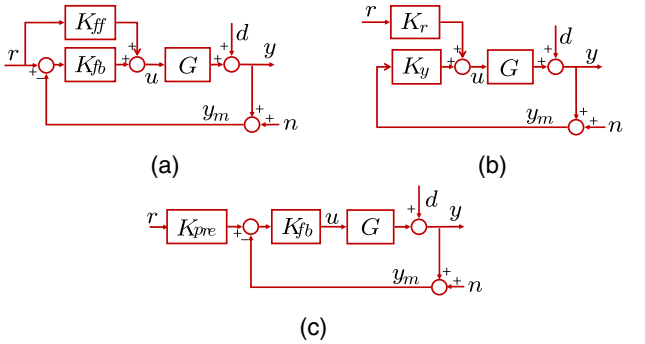

Figure 4. 2DOF control architectures: (a) the feedforward–feedback scheme where the actuation signal $u = K_{\rm ff}r + K_{\rm fb}(r - y_{\rm m})$, (b) Another scheme where $u = K_r r + K_y y_{\rm m}$, and (c) prefilter architecture where $u = K_{\rm fb}(K_{\rm pre}r - y_{\rm m})$. The schemes (a) and (b) are equivalent as control designs in that one can be retrieved exactly in terms of the other. Practical implementable designs for the controllers in (a) and (b) can easily be derived from the control design in (c), however the opposite situation may require certain factorization procedures.

this limitation. For instance, small $|S(\mathrm{j}\omega)|$ over a specified bandwidth might not leave out enough frequency range to zero out the area in the 'finite-waterbed effect' even with $S(\mathrm{j}\omega)$ at the allowed peak value for the remaining frequencies (see figure 3).

3. Two-degree-of-freedom control design methods

In this paper, we show that the feasible space of performance specifications, which are constrained by the limitations described above in the feedback-only configuration, can be extended by using a 2DOF design scheme. In contrast to the feedback-only scheme described in section 2, where the controller acts only on the difference between the reference r and the position-measurement $y_{\rm m}$, in the 2DOF scheme, the controller acts independently on them. This scheme is implemented in *equivalent* multiple architectures as shown in figure 4.

The feedback-only control scheme is indeed a special case of the 2DOF scheme, where $K_{\rm ff} = 0$ in (a), $K_r = -K_y$ in (b), and $K_{\rm pre} = 1$ in (c) in figure 4. Therefore, 2DOF control design should perform at least as well as the feedback-only design. In this paper we explore to what extent the performance becomes better. In this following, we mainly explain the design in the feedforward–feedback structure (where $u = K_{\rm ff}r + K_{\rm fb}(r - y)$, (a) in figure 4) for convenience. In the 2DOF scheme, the robustness to modeling uncertainties as well as the resolution of the device are determined only by the feedback part of the controller, that is the transfer function from d to y that characterizes the robustness to modeling uncertainties is still determined by the sensitivity function $S = (1 + GK_{\rm fb})^{-1}$, and the transfer function from n to y that characterizes resolution is still determined by the complementary sensitivity function $T = (1 + GK_{\rm fb})^{-1}GK_{\rm fb}$. The main difference and advantage in 2DOF control design compared to the feedback-only design stems from the fact that the transfer functions from r to y and from n to y are different and can be designed independently. This difference gives greater independence in designing for better trade-offs between different performance objectives. We use T_{yr} and S_{er} to denote the transfer function from r to y and from r to e, respectively, that is $S_{er} = S(1 - GK_{\rm ff})$, $T_{yr} = SG(K_{\rm ff} + K_{\rm fb})$. In this notation, the relevant closed-loop signals are given by

$$\begin{aligned} &\text{position:} && y = T_{yr}r - Tn + Sd, \\ &\text{tracking error:} && e = S_{er}r + Tn - Sd, \\ &\text{control (actuation) signal:} && u = S(K_{\rm ff} + K_{\rm fb})r - SK_{\rm fb}n - SK_{\rm fb}d. \end{aligned} \tag{5}$$

The control objectives translate to small roll-off frequency as well as high roll-off rates for T to have good resolution, a long range of frequencies for which S_{er} is small to achieve large bandwidth, and low (near 1) values of the peak in the magnitude plot of $S(\mathrm{j}\omega)$ for robustness to modeling uncertainties. Even though the 2DOF control design has greater flexibility than the feedback-only design, the main challenges to design still arise from practical and algebraic (albeit fewer) limitations. The constraints on hardware implementation in terms of sampling frequencies as well as saturation limits of actuation signals limit the scope of this design too. Similarly, the algebraic limitations constrain the control design in this setting too; for instance, the constraint $S(\mathrm{j}\omega) + T(\mathrm{j}\omega) \equiv 1$ and the Bode integral law have the same ramifications on the trade-off between the resolution and robustness to modeling uncertainties as in the feedback-only design.

Optimal control design

The design of control laws for achieving *simultaneously* multiple performance objectives under multiple limitations discussed above makes tuning-based control designs (such as proportional–integral–derivative (PID)) impractical and

unviable. In this paper, we use a framework based on robust control theory, where it is first determined if a set of design specifications is feasible, and when feasible the control law K is obtained by posing and solving an optimization problem. The main advantage of using this optimization framework is that it incorporates performance objectives directly into its cost function. This eliminates the tedious task of tuning gains (in a trial-and-hit manner) as in the PID designs (where even the exhaustively tuned gains may fail to yield acceptable performance). These optimization problems are of the form

$$\min_{K} \|\Phi(K)\|, \tag{6}$$

where Φ is a matrix transfer function whose elements are in terms of the closed-loop transfer functions in (5) and $\|(\cdot)\|$ represents a metric on the transfer functions. Here we use the $\mathcal{H}_\infty$ norm as a metric on the transfer functions given by

$$\|\Phi\|_\infty = \sup_{\omega} \sigma_{\max}(\Phi(\mathrm{j}\omega)), \tag{7}$$

where $\sigma_{\max}$ represents the maximum singular value of the matrix. We incorporate the design objectives into the formulation of the transfer function Φ in two ways. (1) We first specify a target transfer function T_{ref} that guarantees the design specifications and $\Phi(K) = T_{\text{ref}} - T_{\text{cl}}$ represents the mismatch between the closed-loop device transfer function T_{cl} and the target transfer function T_{ref}. In this formulation small $\|\Phi\|$ implies that the mismatch error between the actual closed-loop device and target device is small. (2) We interpret the design specifications in terms of closed-loop signals z (called regulated variables, such as the tracking error e in (5)) and then set Φ as the transfer function from external variables w (called exogenous variables, such as the reference signal r) to these regulated variables. In this formulation, making $\|\Phi\|$ small implies that the ratio of magnitudes of regulated variables to external variables is small regardless of the external signals. In fact, minimization in the $\mathcal{H}_\infty$ norm is strict in the sense that it minimizes the magnitude of z for the *worst* case of exogenous variables w.

Remark. It should be noted that these optimization problems have been studied extensively and there exist standard software routines (for instance in Matlab) where $\mathcal{H}_\infty$-norm problems can be solved. □

In this section, we present three 2DOF control designs motivated by three practical scenarios. In the first case, we assume that the feedback controller is given and we need to design an *optimal* feedforward controller that aims at increasing the bandwidth over the existing feedback-only design. In the second case, we pose an optimal control problem where we solve simultaneously for both the feedforward and feedback controllers that achieve certain performance specifications. In the third case, we again solve for the feedforward and feedback components that wrap around an existing feedback-only design in order to give better performance with extra emphasis on improving the robustness to modeling uncertainties.

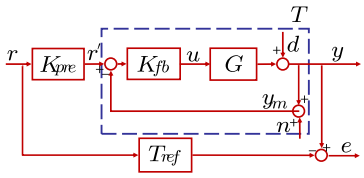

Figure 5. Model matching through the prefilter problem.

3.1. Optimal prefilter model matching design

In some positioning systems, there is a pre-designed feedback component K_{fb} which cannot be replaced or changed (for instance, some commercial scanners come with feedback components designed to accomplish specific tasks such as raster scanning). However, typically there are no such restrictions on the feedforward control design since it can be easily implemented as a prefilter on the reference signal. In the design presented here, the feedforward component K_{pre} is so chosen so that the closed-loop positioning system mimics a *target* transfer function T_{ref} (figure 5). This target transfer function T_{ref} is chosen so that it satisfies desired performance objectives. An advantage of using such model matching schemes is that desired transient characteristics (such as settling times and overshoots) can be incorporated by choosing the appropriate model T_{ref}, and since the closed-loop device is designed to mimic the model, it inherits the transient characteristics too. After noting that the closed-loop device transfer function from r to y is given by $T K_{\text{pre}}$, the feedforward component K_{pre} is chosen by solving an optimization problem such that the $\mathcal{H}_\infty$-norm of the mismatch transfer function $E = T_{\text{ref}} - T K_{\text{pre}}$ is minimized. Small values of $\|E\|_\infty$ guarantee small values for the mismatch error signal (see figure 5) given by

$$e = T_{\text{ref}} r - y = (T_{\text{ref}} - T K_{\text{pre}}) r. \tag{8}$$

To ensure practical implementation, it is assumed that T_{ref} and T are stable, proper transfer functions.

Note that this optimization problem is trivial if T is a minimum phase transfer function, that is if it has only stable zeros. In this case T^{-1} is stable and the solution K_{pre} can be easily obtained as $T^{-1} T_{\text{ref}}$. However, typical nanopositioning systems are flexure based with non-collocated actuators and sensors, which typically manifest as non-minimum phase zeros of T. In this case, the optimal solution can be found by applying Nevanlinna–Pick theory [22] as follows.

The model matching problem is equivalent to finding the minimum γ such that $\|T_{\text{ref}} - T K_{\text{pre}}\|_\infty \leqslant \gamma$, where the minimum $\gamma = \gamma_{\text{opt}}$ is achieved for some stable K_{pre}. If we define $E_\gamma = \frac{1}{\gamma}(T_{\text{ref}} - T K_{\text{pre}})$ for $\gamma > 0$, then this problem can be restated as finding a stable K_{pre}, which requires $\|E_\gamma\|_\infty \leqslant 1$. Note that, for stable K_{pre}, E_γ satisfies the interpolating conditions $E_\gamma(z_i) = \frac{1}{\gamma} T_{\text{ref}}(z_i)$ for every non-minimum phase zero z_i of the scanner G. Therefore we can cast this as a Nevanlinna–Pick (NP) problem, of finding a function E_γ in the space of stable, complex-rational functions such that it interpolates

$$\{(z_i, E_\gamma(z_i))\}_{i=1}^{n}. \tag{9}$$

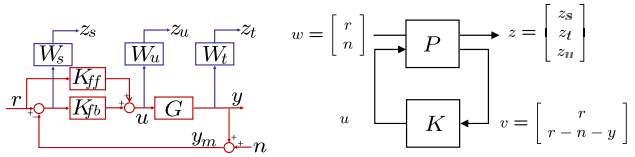

Figure 6. Mixed sensitivity problem for 2DOF control design. The signals z_s, z_t, and z_u represent tracking error, the noise component in the position signal, and the control signal. The weights W_s, W_t, and W_u are chosen to reflect the design specifications of tracking bandwidth, positioning resolution, and saturation limits on the control signal. To achieve these objectives, a control design $K = [K_{\rm ff}\ K_{\rm fb}]$ which minimizes the $\mathcal{H}_\infty$-norm of the transfer function from $w = [r^{\rm T}\ n^{\rm T}]^{\rm T}$ to $z = [z_s^{\rm T}\ z_t^{\rm T}\ z_u^{\rm T}]^{\rm T}$ is sought through the optimal control problem. In typical software that solve this optimization problem, the open-loop matrix transfer function P that relates $[w^{\rm T}\ u^{\rm T}]^{\rm T}$ to $[z^{\rm T}\ v^{\rm T}]^{\rm T}$ serves as the input parameter to the optimization routine. In the 2DOF framework the controller has access to both the reference r and the measured error $r - n - y$ signals.

Moreover, it can be shown that $\gamma_{\rm opt}$ is equal to the square root of the largest eigenvalue of the matrix $A^{-\frac{1}{2}} B A^{-\frac{1}{2}}$, where the elements of matrix A and B are, respectively [22]

$$a_{ij} = \frac{1}{z_i + \bar{z_j}}, \qquad b_{ij} = \frac{b_i \bar{b_j}}{z_i + \bar{z_j}}. \tag{10}$$

The prefilter is given by

$$K_{\rm pre} = T^{-1}(T_{\rm ref} - \gamma_{\rm opt} E_\gamma). \tag{11}$$

Remark. This method requires additional steps for better performance. First, for better tracking, $K_{\rm pre}$ need to be scaled so that the closed-loop function matches the target transfer function at 0 Hz (i.e. the dc gain match). Second, since the obtained $K_{\rm pre}$ may not be proper, a weight function W_0 which is a low pass filter needs to be multiplied so that the $W_0 K_{\rm pre}$ becomes proper. □

3.2. 2DOF mixed sensitivity synthesis

In this control synthesis scheme, both the feedforward and the feedback control laws are solved in an optimal control setting. In order to reflect the performance objectives and physical constraints, the regulated outputs were chosen to be the weighted tracking error, $z_s = W_s e$, the weighted system output, $z_t = W_t y$, and the weighted control input, $z = W_u u$ (figure 6).

Using (5), the closed-loop matrix transfer function from the exogenous variables $w = [r\ n]^{\rm T}$ to the regulated variables $z = [z_s\ z_t\ z_u]$ is given by

$$\begin{bmatrix} z_s \\ z_t \\ z_u \end{bmatrix} = \underbrace{\begin{bmatrix} W_s S_{er} & -W_s S \\ W_t T_{yr} & -W_t T \\ W_u S(K_{\rm ff} + K_{\rm fb}) & -W_u S K_{\rm fb} \end{bmatrix}}_{=\Phi(K)} \begin{bmatrix} r \\ n \end{bmatrix}. \tag{12}$$

Accordingly, the $\mathcal{H}_\infty$ optimal control problem that we pose is $\min_K \|\Phi(K)\|_\infty$. The minimization of z_s reflects the tracking-bandwidth requirement. If we design the *weight function* $W_s(\mathrm{j}\omega)$ to be large in a frequency range $[0\ \omega_{\rm BW}]$ and ensure that z_s is small over the entire frequency range (through the above optimization problem) then the tracking error e will be small in the frequency range $[0\ \omega_{\rm BW}]$; that is, the closed-loop positioning device has a bandwidth $\omega_{\rm BW}$. Alternatively, note that the transfer function from r to z_s is $W_s S_{er}$. The optimization problem along with our choice of W_s will ensure that the transfer function S_{er} is small in the frequency range $[0\ \omega_{\rm BW}]$. Similarly, the transfer function from n to z_t is the weighted complementary sensitivity function $W_t T$, whose minimization ensures better resolution as it forces low control gains at high frequencies, and the transfer function from r to z_u is $W_u S(K_{\rm ff} + K_{\rm fb})$, which measures the control effort. Its minimization reflects in imposing the practical limitation of the control signals to be within saturation limits.

Remark. The choice of *weight functions* is crucial in this design. The issue of conflicting requirements on the design of S and T transfer functions arising from competing performance objectives is resolved by achieving compromises (trade-offs) between them by designing appropriate weight transfer functions. The weight functions are used to scale closed-loop transfer functions to specify the frequency information of the performance objectives and system limitations. For instance, including $\|W_t T\|_\infty$ instead of $\|T\|_\infty$ to be small (for achieving resolution objective) in the optimization problem, where the weight function $W_t(\mathrm{j}\omega)$ is designed as a high-pass filter, forces the control solution such that $T(\mathrm{j}\omega)$ is small at high frequencies (and not necessarily at low frequencies). This weight function allows a compromise, for instance, between the resolution and bandwidth requirements. More specifically, if the condition for $\|T\|_\infty$ to be small is imposed instead (i.e. $|T(\mathrm{j}\omega)|$ over all frequencies), the limitation from the Bode integral law leaves little or no room for achieving the robust performance objective. □

In (12), weight functions $W_s(\mathrm{j}\omega)$ and $W_t(\mathrm{j}\omega)$ are to be so designed that the optimal control problem simultaneously enforces S_{er} to be small in a frequency range that will determine the bandwidth of the positioning system, and the peak value $\|S\|_\infty$ to be small to ensure robustness to modeling uncertainties, as well as T to roll-off at high frequencies to provide better resolution. Note that the optimization problem is severely constrained since we have only these *two* weight functions to shape *four* transfer functions S_{er}, T_{yr}, S, and T, which in turn are limited by the algebraic constraints such as $S_{er} + T_{yr} = 1$, $S + T = 1$. The design of the weight function W_u is relatively simpler as it has only one objective of control saturation to account for. In order to alleviate this problem, we introduce two new weight functions W_r and W_n that carry the information content of input signals r and n, respectively. Thereby, the transfer matrix for optimization in (12) is replaced by

$$\begin{bmatrix} z_s \\ z_t \\ z_u \end{bmatrix} = \begin{bmatrix} W_s S_{er} W_r & -W_s S W_n \\ W_t T_{yr} W_r & -W_t T W_n \\ W_u S(K_{\rm ff} + K_{\rm fb}) W_r & -W_u K_{\rm fb} S W_n \end{bmatrix} \begin{bmatrix} r \\ n \end{bmatrix}. \tag{13}$$

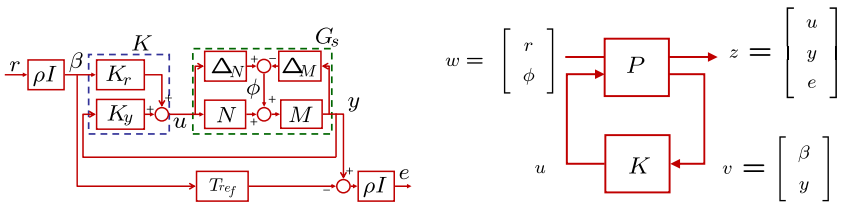

Figure 7. Control design for model matching and robustness to modeling errors. The signals u, y, and e that represent the control signal, the noise component in the position signal, and the mismatch error are chosen as regulated variables z. The transfer function from the reference signal r to these regulated variables z reflects the performance objectives of bounded control signals and noise attenuation, as well as the model matching. The transfer function from the effects of modeling error ϕ to z represents the effect of modeling errors (unmodeled dynamics) on performance objectives. To achieve robust performance, a control design $K = [K_r\ K_y]$ which minimizes the $\mathcal{H}_\infty$-norm of the transfer function from $w = [r^{\mathrm{T}}\ n^{\mathrm{T}}]^{\mathrm{T}}$ to $z = [u^{\mathrm{T}}\ y^{\mathrm{T}}\ e^{\mathrm{T}}]^{\mathrm{T}}$ is sought through the optimal control problem.

Remark. Note that the inclusion of weights W_r and W_n does alleviate, to an extent, the severity of constraints on the optimal control problem which requires shaping *four* transfer functions. Even with these additional weighting functions, the shaping of these transfer functions is still not unconstrained since they are not independent even though the input-weighting functions (W_r, W_n) and the output-weighting functions (W_s, W_t) can be chosen independently. The limitation stems from these weight functions appearing as products in (13), which reduces the flexibility in their choice. □

3.3. 2DOF optimal robust model matching designs

In some nanopositioning systems, the issue of robustness to modeling uncertainties is critical. Systems with pre-designed feedback control have satisfactory performance (resolution and bandwidth) when experiments are conducted very carefully at 'near nominal' operating conditions. However, there is rapid degradation of performance (it sometimes even become unstable) when the operating conditions deviate from the nominal condition. In [23], control designs based on the Glover–McFarlane method [24, 25] that wrapped around pre-existing controllers were implemented that resulted in substantial improvements in robustness. Here we present a control design where we pose and solve an optimal control problem that simultaneously designs a wrap-around feedback controller for robustness as well as the feedforward controller for better bandwidth. We use a 2DOF control design developed in [26, 27] that uses model matching where the transfer function of the closed-loop device is made to match a pre-specified target transfer function T_{ref}.

In this design formulation, we represent the modified scanner system by $G_s = GK_s$, where G represents the scanner and K_s represents the pre-existing controller. The corresponding 2DOF optimal control problem can be explained in terms of figure 7. The optimization routine seeks $K = [K_r\ K_y]$ such that the closed-loop system guarantees 'optimal' robustness to modeling uncertainties as well as minimizing the mismatch between the transfer function from r to y and a target transfer function T_{ref}. The condition for robustness is imposed by requiring the controller to guarantee stability for a set of transfer function models that are 'close' to the nominal model G_s. More specifically, the optimal control problem is cast in such a way that the resulting controller guarantees the stability of the closed-loop positioning system where the shaped scanner system is represented by *any* transfer function G_p in the following set:

$$\{G_p | G_p = (M - \Delta_M)^{-1}(N + \Delta_N), \quad \text{where } \|[\Delta_M\ \Delta_N]\|_\infty \leqslant \gamma^{-1}\}, \tag{14}$$

where $G_s = M^{-1}N$ is a coprime factorization [28], $[\Delta_M\ \Delta_N]$ represents the uncertain (unmodeled) dynamics, and γ specifies a bound on this uncertainty.

Remark. This characterization of robustness is specifically apt for many nanopositioning systems which typically have very low damping coefficients (<0.01 are common). Since the set described by G_p can have marginally stable/unstable systems, the critical robustness issue is adequately incorporated in this design. The model matching problem is incorporated by seeking the control design K such that the mismatch between the closed-loop transfer function and T_{ref} is minimized. The desired closed-loop transfer function T_{ref} is selected to satisfy the desired response characteristics. The parameter ρ determines the emphasis between model matching and robustification in optimization. (If ρ is 0 then the objective becomes the same as the Glover–McFarlane design problem.) □

The optimal control problem is cast in the $\mathcal{H}_\infty$ optimal control set-up (described in the previous section). More specifically, the signals $z = [u^{\mathrm{T}}\ y^{\mathrm{T}}\ e^{\mathrm{T}}]^{\mathrm{T}}$ are chosen as regulated variables and the controller K is obtained which minimizes the $\mathcal{H}_\infty$ norm of the transfer function Φ_{zw} from $w = [r^{\mathrm{T}}\ \phi^{\mathrm{T}}]^{\mathrm{T}}$ to z (see figure 7) described by

$$\begin{bmatrix} u \\ y \\ e \end{bmatrix} = \underbrace{\left[\begin{array}{c|c} \rho K_r S & K_y S M^{-1} \\ \rho G_s K_r S & S M^{-1} \\ \hline \rho^2 (G_s K_r S - M_0) & \rho S M^{-1} \end{array}\right]}_{=\Phi_{zw}} \begin{bmatrix} r \\ \phi \end{bmatrix}, \tag{15}$$

where $S = (1 - G_s K_y)^{-1}$. Here the exogenous signal ϕ represents a disturbance signal that is the effect of the unmodeled dynamics. Thus by making the norm on the matrix transfer function small, we make the closed-loop device insensitive to the effects of unmodeled dynamics. For instance, the smaller the $\|\Phi_{zw}\|_\infty$, the smaller is the effect of the disturbance ϕ on the mismatch signal e. This implies that the mismatch signal has been made insensitive to the

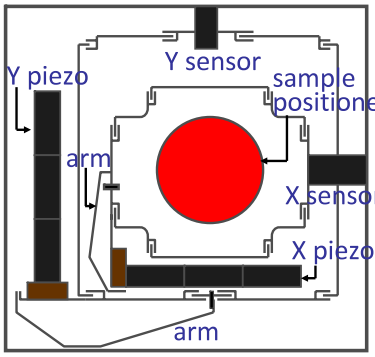

Figure 8. Schematic of flexure scanner. The sample is placed on the central block of the flexure stage which is driven by the x piezo which in turn is driven by the y piezo. The x and y sensors measure the stage position.

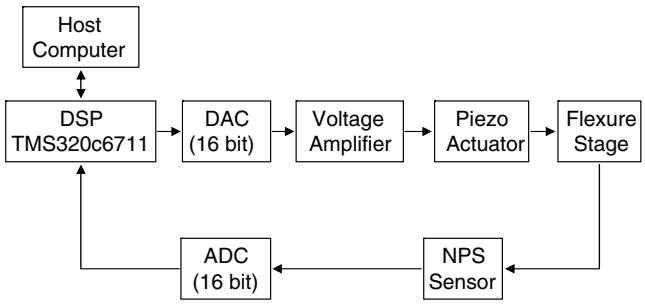

Figure 9. Control system set-up for the flexure stage.

modeling errors. The signals u and y are incorporated as regulated variables in the optimization problem to account for the control-saturation constraint and for the noise attenuation, respectively.

Remark. An additional step is required to improve the tracking performance. As a final refinement for the tracking problem, K_r needs to be scaled so that the closed-loop transfer function matches the reference transfer function for the steady state problem as in the optimal prefilter model matching design. A scale constant W_0 is defined as $W_0 = S(s)[G_s(s)K_y(s)]^{-1}T_{\text{ref}}|_{s=0}$, and the resulting controller becomes $K = [K_r W_0 \; K_y]$. □

4. Experimental implementation and demonstration of control designs

This section demonstrates the optimal control framework presented above on a two-dimensional flexure scanner of molecular force probe (MFP-3D) from Asylum Research Inc., Santa Barbara, CA.

4.1. Device description

A schematic of the nanopositioning system (the scanner) is shown in figure 8. It has two flexure components with component 'X' stacked over the 'Y' where the sample-holder is carried by the 'X'-component. Each stage, by virtue of the serpentine spring design, deforms under the application of force, providing motion. These forces are generated by stack-piezos. There are three piezoactuators in series for each axis. The motion of each flexure component is measured by the corresponding nanopositioning sensors which is modified from the linear variable differential transformer (LVDT) and the associated demodulation circuit. The piezoactuators lead to a travel range of 90 μm in a closed loop in both directions. The nanopositioning sensors have noise less than 0.6 nm (deviation) over 0.1–1 kHz bandwidth.

The control law is discretized and implemented on a Texas Instrument TMS320C6713 DSP using code composer studio with 16-bit A/D and 16-bit D/A channels. The set-up of the control system is shown in figure 9.

4.2. Identification

Physical modeling of the device is difficult due to its complicated structural design and poorly understood piezoactuation phenomena, and attempts in that direction typically result in complex dynamic models which are not easy to design control for. Therefore, identification techniques were used to derive linear models about an operating point, where the sensor output gave a 'zero' reading corresponding to 'zero' input to the piezoactuators. The device is viewed as a two-input–two-output system in which the low-voltage signals to the X and Y amplifiers are the inputs and the motions of X and Y flexure-stage components measured by the corresponding sensors are the outputs. This results in four input–output transfer functions G_{ij}, i, j in $\{x, y\}$. Here G_{ij} represents the transfer function from input j to output i. The frequency-response-based identification was done where a sine-sweep over a bandwidth from 1 Hz to 2 kHz with amplitude 10 mV was given to each axis using an HP 35670A signal analyzer. From the identification results, both the X and the Y crosstalk, represented by G_{xy} and G_{yx}, are seen to be relatively small ($\|G_{xy}(\mathrm{j}\omega)\|_\infty$ and $\|G_{yx}(\mathrm{j}\omega)\|_\infty$ are less than -17.76 dB), which is expected since, by design, the X and Y flexure components are decoupled and are orthogonal to each other. Therefore, the nanopositioning system is modeled by two independent single-input–single-output (SISO) units. The mode of operation of this device is such that higher-bandwidth requirements are made on the smaller stage X whereas the Y stage is made to move relatively slowly. Hence, there is a greater emphasis on the control designs for the X stage, which is presented in this paper.

This process was repeated to obtain the frequency responses of the system at different operating points (by giving various dc offsets) spanning the range of operation of the device (figure 10(a)). The variation in these responses is indicative of the modeling errors (uncertainties) in our identification scheme. In addition, it was observed that the frequency response at the same operating point varies when obtained at different times. In view of these uncertainties, robustness of the closed-loop system is a critical requirement of control design. The nominal frequency response of the system is obtained from averaging five experiments on the nominal

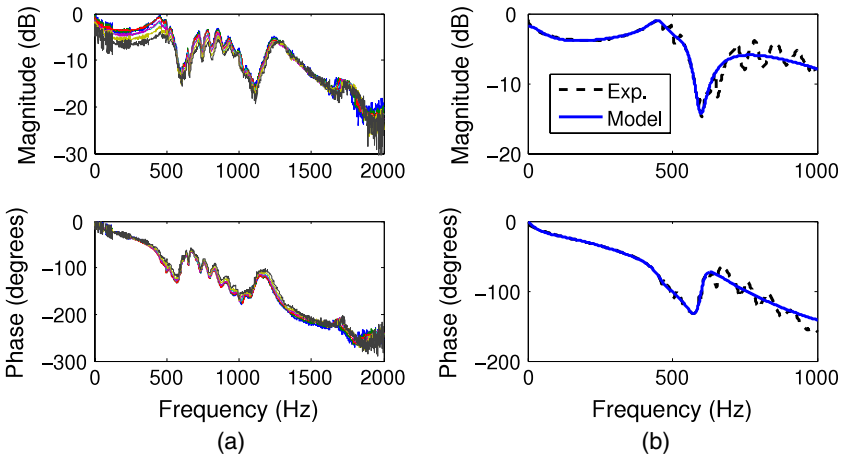

Figure 10. (a) Experimental frequency responses at various operating positions. (b) Nominal frequency response(dashed) and model frequency response (solid).

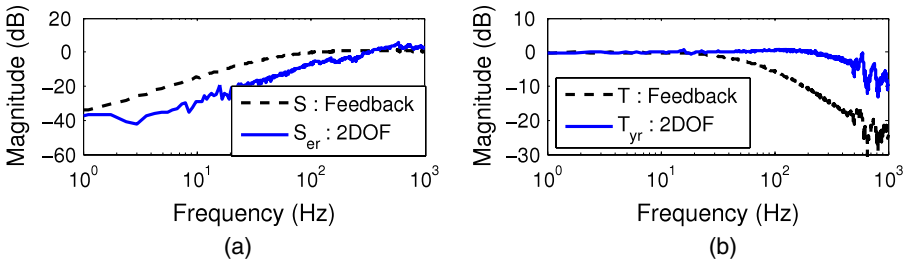

Figure 11. Comparison of experimentally obtained magnitude of $S(s)$ and $T(s)$ from an $\mathcal{H}_\infty$ feedback-only control design (dashed) with $S_{er}(s)$ and $T_{yr}(s)$ from a prefilter model matching 2DOF control (solid). The feedforward controller designed using the prefilter model matching design achieves over 330% improvement in the tracking bandwidth of the closed-loop design. The robustness and resolution are determined by the feedback components S and T, and therefore remain the same for the two cases.

operating point which is at dc offset corresponding to 0 V output value. Figure 10(b) shows the Bode diagram of the fitted mathematical model with the nominal experimental result. Weighted iterative least square fitting was performed over 0–1 kHz and the reduction through balanced realization [29] resulted in the following seventh-order model:

$$G_{xx}(s) = \frac{-1122.3157(s - 1.152 \times 10^4)(s + 543)}{(s + 390.3)(s^2 + 470.9s + 8.352 \times 10^6)} \times \frac{(s^2 + 587.2s + 8.628 \times 10^6)}{(s^2 + 689.4s + 1.315 \times 10^7)} \times \frac{(s^2 + 226.5s + 1.407 \times 10^7)}{(s^2 + 4950s + 2.44 \times 10^7)}. \quad (16)$$

This seventh-order model did not capture the dynamics beyond 500 Hz, as shown in figure 10(b). Its use is justified since the frequency range of interest is less than 500 Hz and larger models result in implementations of higher-order control which cannot be accommodated by the processor with short sampling time. This modeling uncertainty from using a low-order model was accounted for by imposing the requirement of making the closed-loop system robust to it in the control design.

4.3. Control design, implementation and analysis

The three types of 2DOF control design described in section 3 were applied to the MFP-3D scanner.

4.3.1. Optimal prefilter model matching. For the purpose of illustration, we designed the feedback 'pre-existing controller' by using a feedback-only design with an $\mathcal{H}_\infty$ optimal framework which achieves much better bandwidth when compared to PI/PID designs [11]. Therefore, improvements resulting from the application of our model matching shown in this paper are even more significant when applied to typical scanners that commonly use PI/PID-based feedback controllers. A ninth-order feedback controller K_{fb} was obtained as a result of a mixed sensitivity $\mathcal{H}_\infty$ optimization for feedback-only designs given by

$$K_{\text{fb}} = \frac{872\,851.2498(s + 3.142 \times 10^5)(s + 390.3)}{(s + 1.123 \times 10^7)(s + 7.588 \times 10^4)} \times \frac{(s^2 + 470.9s + 8.352 \times 10^6)}{(s + 1.261 \times 10^4)(s + 543.5)(s + 0.031\,42)} \times \frac{(s^2 + 689.4s + 1.315 \times 10^7)}{(s^2 + 587.1s + 8.628 \times 10^6)} \times \frac{(s^2 + 4950s + 2.44 \times 10^7)}{(s^2 + 226.7s + 1.407 \times 10^7)} \quad (17)$$

This feedback-only design yielded a bandwidth of 49.4 Hz, a roll-off frequency of 60.1 Hz, and $\|S\|_\infty$ of 1.15 for the closed-loop device (represented by solid lines in figure 11).

The prefilter is designed in pursuit of matching T as $T_{\text{ref}} = \frac{1}{0.0003s+1}$ by using the Nevanlinna–Pick solution, which

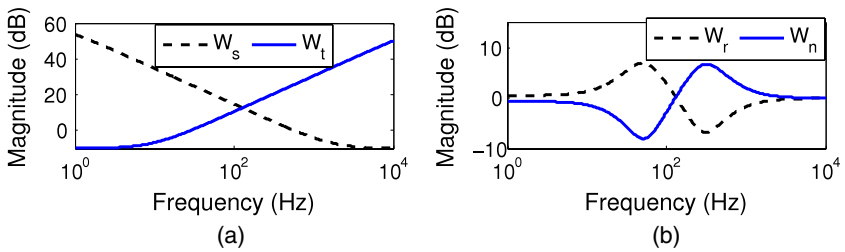

Figure 12. Choice of weight functions.

gives the control law as the following:

$$K_{\text{pre}} = \frac{2.0289 \times 10^{-10}(s + 1.123 \times 10^7)(s + 7.626 \times 10^4)}{(s + 3.142 \times 10^5)} \times \frac{(s + 1.18 \times 10^4)(s + 544.6)(s + 349.9)}{(s + 2857)(s + 543)}. \quad (18)$$

Before directly applying this control law, it is slightly modified (using steady state gain matching and multiplication of the weighting function) as discussed in section 3.1. The control law from the NP solution is improper and has relative degree equal to -2, and therefore we multiply it by the weight function $W_0 = \frac{1}{(1\times10^{-4}s+1)^2}$ so that it becomes proper. The scaling factor 1.25 also multiplies $W_0 K_{\text{pre}}$ for the dc gain match.

Figure 11 shows the experimentally obtained transfer function from the reference to the error, i.e. S from the feedback control and S_{er} from the 2DOF control, which represent the tracking performance ($\omega_{\text{BW}} = 49.4$ Hz (feedback), $= 214.5$ Hz (2DOF)), and the transfer function from the reference to the output, i.e. T of the feedback control and T_{yr} of the 2DOF control. Thus this 2DOF design yields an improvement of over 330% in bandwidth over the feedback-only design. Since the feedback components of the two designs are the same and this component completely determines the robustness to modeling errors (characterized by $\|S\|_\infty$) as well as the positioning resolution (characterized by ω_T), the resolution and robustness remain the same for both devices. An interesting observation is that the tracking bandwidth ω_{BW} is greater than the roll-off frequency $\omega_T = 60.1$ Hz, which is impossible in feedback-only design.

4.3.2. 2DOF mixed sensitivity synthesis. As discussed in section 3, four weight functions (W_r, W_n, W_s, and W_t) are chosen to shape closed transfer functions: S_{er} with $W_s W_r$, S with $W_s W_n$, T_{yr} with $W_t W_r$, and T with $W_t W_n$. The performance objectives of high bandwidth, high resolution, and robustness to modeling errors were reflected as follows. High resolution requires the roll-off frequency of ω_T to be small, that is T to be small beyond ω_T. This is imposed by designing the weight frequency $W_t = \frac{1000s+5.961\times10^4}{s+1.885\times10^5}$ to be large at high frequencies (we chose ω_T around 75 Hz). The range of frequencies where S is small (required for small tracking error) is restricted to small frequencies since small T at high frequencies implies S to be near 1 at high frequencies. Thus $W_s = \frac{0.3162s+3456}{s+3.456}$ ensures S is small at low frequencies and allows for its cross-over frequency to be small enough to make the designed ω_T feasible. Figure 12 shows the choice of the weight functions that reflect these objectives. The choice of W_r and W_n is made such that at the frequency the W_t starts increasing, W_r starts increasing and W_n starts decreasing (in fact W_n is chosen as the inverse of W_r). Note that at the frequency the W_s stop rolls off, W_r and W_n converge to 1. This choice of W_r and W_n ensures that S_{er} is small even when S is not small. This is done by exploiting that S is shaped by W_s while S_{er} is shaped by $W_s W_r$. The choice of weight function $W_u = 0.1$ restricted the control signal values to be within saturation limits.

The result of $\mathcal{H}_\infty$ synthesis yielded the following control laws, and the corresponding closed-loop transfer function are shown in figure 13.

$$K_{\text{ff}} = \frac{7.520 \times 10^9(s + 1.885 \times 10^5)(s + 1.288 \times 10^4)}{(s + 1.122 \times 10^9)(s + 3.041 \times 10^6)(s + 1.3 \times 10^4)} \times \frac{(s + 2623)(s + 723.3)(s + 390.3)}{(s + 3725)(s + 1001)(s + 543)} \times \frac{(s + 275.9)(s + 37.31)(s + 3.413)}{(s + 171)(s + 144)(s + 3.412)} \times \frac{(s^2 + 470.9s + 8.352 \times 10^6)}{(s^2 + 1965s + 3.426 \times 10^6)} \times \frac{(s^2 + 689.4s + 1.315 \times 10^7)}{(s^2 + 587.2s + 8.628 \times 10^6)} \times \frac{(s^2 + 4950s + 2.44 \times 10^7)}{(s^2 + 226.5s + 1.407 \times 10^7)}$$

$$K_{\text{fb}} = \frac{9.008 \times 10^9(s + 1.885 \times 10^5)(s + 496.9)}{(s + 1.122 \times 10^9)(s + 3.041 \times 10^6)(s + 1.3 \times 10^4)} \times \frac{(s + 390.3)(s^2 + 1341s + 1.691 \times 10^6)}{(s + 3725)(s + 1001)(s + 543)} \times \frac{(s + 144)(s^2 + 1965s + 3.426 \times 10^6)}{(s + 171)(s + 144)(s + 3.412)} \times \frac{(s^2 + 470.9s + 8.352 \times 10^6)}{(s^2 + 1965s + 3.426 \times 10^6)} \times \frac{(s^2 + 689.4s + 1.315 \times 10^7)}{(s^2 + 587.2s + 8.628 \times 10^6)} \times \frac{(s^2 + 4950s + 2.44 \times 10^7)}{(s^2 + 226.5s + 1.407 \times 10^7)}. \quad (19)$$

The feedforward and feedback control laws obtained from $\mathcal{H}_\infty$ mixed sensitivity synthesis procedure were implemented.

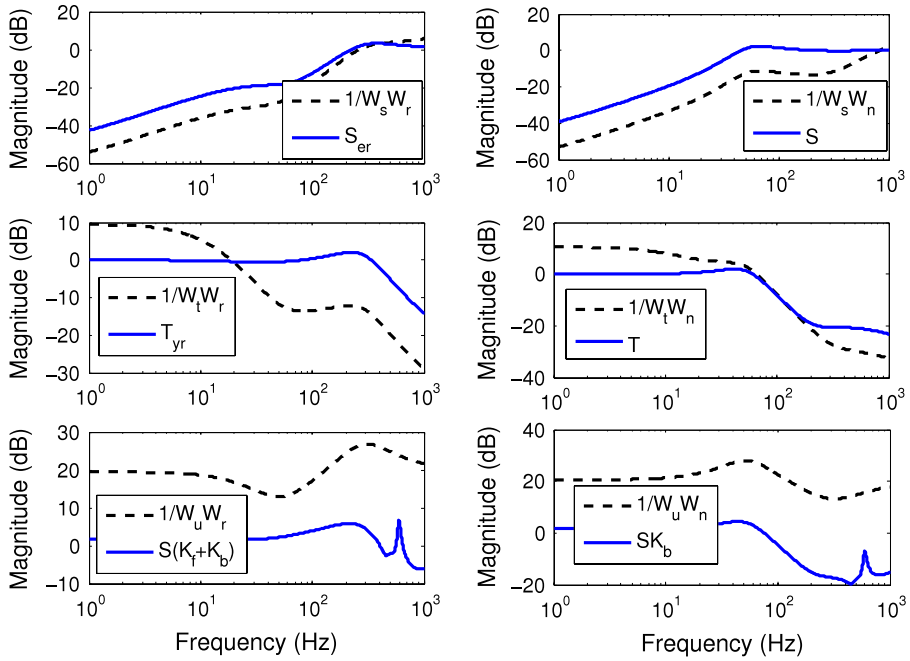

Figure 13. Mixed sensitivity synthesis of 2DOF control.

Figure 14(a) shows the experimentally obtained transfer function from the reference to the error, i.e. $S_{er}(s)$, which represents the tracking performance ($\omega_{\mathrm{BW}} = 148.2$ Hz), and the transfer function from the disturbance to the error, i.e. $S(s)$, and figure 14(b) shows the transfer function from the reference to the output, i.e. $T_{yr}(s)$, and the transfer function from the noise to the output, i.e. $T(s)$. Comparing with the roll-off frequency $\omega_T = 75.4$ Hz, the tracking bandwidth ω_{BW} is higher, as in optimal prefilter model matching design. There was in improvement of 290% in bandwidth for the same values of resolution and robustness if compared to the feedback-only design. Similarly, an improvement in other performance objectives (resolution and robustness) can be obtained by appropriately designing the weight functions (not presented in this paper).

4.3.3. 2DOF optimal robust model matching. Proportional–integral (PI) and proportional–integral–integral (PII) controllers are the most common controllers currently used in commercial scanning-probe microscopes. Their popularity stems from the fact that they are simple to implement, and easy for users to develop a feel for tuning them. Moreover, PII controllers track ramp signals, which can partially represent raster scan, with zero steady state error. The PII controller has the structure $K_{\mathrm{PII}} = k_p + \frac{k_i}{s} + \frac{k_{ii}}{s^2}$. After an exhaustive search over the space of controller parameters to meet the performance and robustness requirement, $k_p = 0$ was chosen to get the roll-off of 2 for high-frequency noise attenuation and $k_i = 1.1 \times 10^3$ and $k_{ii} = 6.2 \times 10^4$ were chosen to meet the bandwidth and robustness requirements. Experiments with the resulting controller showed the bandwidth of a closed-loop device to be 81 Hz and the roll-off frequency ω_T at 235 Hz, and $\|S\|_\infty$ to be 1.54 (figure 15 shows the experiment results for this PII-based feedback-only design). The design outlined in section 3.3 was applied to G_{xx} with $K_s = K_{\mathrm{PII}}$, $\rho = 3$, $T_{\mathrm{ref}} = \frac{1}{0.0003s+1}$. The 2DOF controller was obtained as

$$\begin{aligned} K_r = {} & \frac{7.762 \times 10^9 (s+56.47)(s+56.36)}{s^2(s+8.286\times10^5)(s+3333)} \\ & \times \frac{(s^2+1116s+3.445\times10^5)}{(s+541.4)(s+56.36)} \\ & \times \frac{(s^2+479.2s+8.373\times10^6)}{(s^2+599.7s+8.553\times10^6)} \\ & \times \frac{(s^2+678.4s+1.318\times10^7)}{(s^2+493.8s+1.458\times10^7)} \\ & \times \frac{(s^2+4962s+2.461\times10^7)}{(s^2+8440s+4.552\times10^7)} \end{aligned} \tag{20}$$

$$\begin{aligned} K_y = {} & \frac{-1.992 \times 10^9 (s+453.8)(s+56.36)}{s^2(s+8.286\times10^5)(s+541.4)} \\ & \times \frac{(s+52.33)(s^2+454.9s+8.264\times10^6)}{(s+56.36)(s^2+599.7s+8.553\times10^6)} \\ & \times \frac{(s^2+713.9s+1.305\times10^7)}{(s^2+493.8s+1.458\times10^7)} \\ & \times \frac{(s^2+4920s+2.411\times10^7)}{(s^2+8440s+4.552\times10^7)}. \end{aligned}$$

Figure 15(a) compares the experimentally obtained response from the reference to the error, i.e. S from the feedback-only control and S_{er} from the 2DOF control, which represent the tracking performance. There is over 64% improvement in the tracking bandwidth ($\omega_{\mathrm{BW}} = 81$ Hz (feedback), $= 133$ Hz (2DOF)). Figure 15(b) compares the response from the reference to the output, i.e. $T(s)$ from the feedback control and $T_{yr}(s)$ from the 2DOF control.

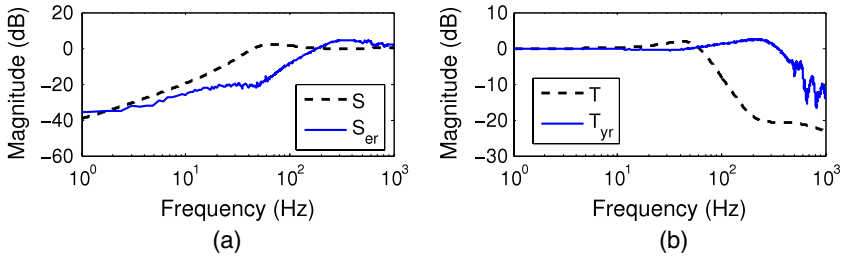

Figure 14. Magnitude of $S_{er}(s)$ and $T_{yr}(s)$ (solid) obtained from experiment and compared to S and T (dashed).

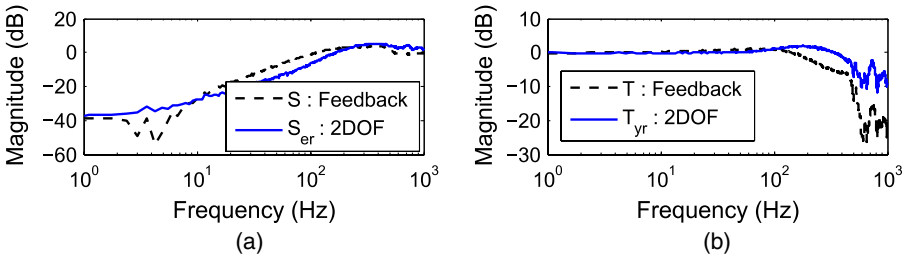

Figure 15. Comparison of experimentally obtained magnitude of $S(s)$ and $T(s)$ from a PII feedback-only control design (dashed) with $S_{er}(s)$ and $T_{yr}(s)$ from a 2DOF optimal robust model matching control (solid). The 2DOF optimal robust model matching controller achieves over 64% improvement in the tracking bandwidth.

Comparison of $S(\mathrm{j}\omega)$ shows the improvement of robustness ($\|S\|_\infty = 1.52$ (feedback) $= 1.21$ (2DOF)). Note that this design process improves the bandwidth and robustness at the same time for fixed resolution, which is generally impossible due to fundamental limitations in feedback-only designs.

As discussed in section 3, this design achieves robustification and model matching simultaneously and wraps around the pre-existing control. Accordingly the closed-loop system with this design achieves better robustness to modeling errors and tracking bandwidth compared to the system with only a pre-existing controller. However, if the pre-existing controller results in a closed-loop device that is robust to modeling errors but has insufficient bandwidth, then this process increases the bandwidth without adversely affecting the robustness, and vice versa [30].

5. Analysis and discussion

5.1. Trade-offs between performance objectives

One of the main advantages of the formulation of these optimal designs, besides providing an apt framework for incorporating design specifications and achieving them, is that their analysis gives significant insights into the trade-offs between difference objectives and relative roles of feedforward and feedback components. For instance, let us analyze the first two rows in equation (12) which reflects the tracking bandwidth and positioning resolution objectives. The term Sn in $z_s = W_s S_{er} r - W_s S n$ and the term Tn in $z_t = W_t T_{yr} r - W_t T n$ quantify the deteriorating effects of noise on tracking bandwidth and positioning resolution, respectively. They cannot be made small in the same frequencies since S and T cannot be made small simultaneously. This problem is resolved by designing weight functions W_s and W_t that separate these objectives in the frequency domain. An inappropriate choice of these weight functions can lead to an infeasible optimization problem.

5.2. Relative roles of feedforward and feedback

Since $S_{er} = S(1 - GK_{\mathrm{ff}})$ and S cannot be made small over the entire bandwidth range (in order to allow for noise attenuation), the feedforward control K_{ff} plays an 'active' role in making S_{er} small beyond the frequency where S is not small (say greater than $1/\sqrt{2}$). Also, since $S = (1 + GK_{\mathrm{fb}})^{-1}$ is completely determined by K_{fb}, the feedback component is dominant in frequencies where S is small. This in turn implies that robustness to modeling errors is solely determined by K_{fb}, and therefore the feedforward component has a very little role in low frequencies (where typically robustness requirements are specified). For instance, figure 16 shows the Bode diagram for $K(s) = [K_r(s)\ K_y(s)]$ in terms of architecture in figure 4(b). Figure 16(a) shows the controller from the optimal prefilter model matching design. Note that, in low frequencies, the control law is essentially feedback-only since K_r and K_y have almost same magnitude and are 180° out of phase, i.e. $K_{tr} \approx -K_{ty}$, and therefore $u \approx K_y(r - y)$. The 2DOF nature of the control becomes active at high frequencies (>50 Hz), that is near the bandwidth achieved by the pre-existing control design. Similarly, in the controller from the 2DOF optimal robust model matching design, the feedforward part is dominant at high frequencies (>100 Hz) (figure 16(b)).

From the above discussion, the main contribution of the feedforward component is in the frequency range where S is no longer small. However, this frequency range is limited. Typically, nanopositioning devices have very poor capabilities to provide motions beyond their flexure resonance frequencies. This is characterized by high roll-off rates of

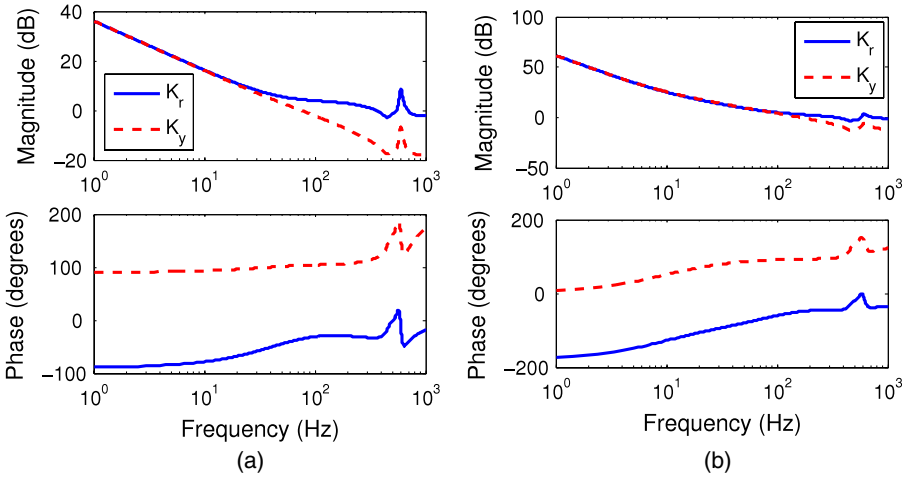

Figure 16. Bode plot $K(s) = [K q_r(s)\ K_y(s)]$: (a) optimal prefilter model matching controller, (b) 2DOF optimal robust model matching controller.

G. Therefore, very high control inputs (high-voltage inputs) are needed to make the positioning systems beyond their flexure resonances. The saturation limits on control signals form the main constraints on attaining bandwidths beyond flexure resonances. Thus the feedforward components provide performance *enhancements* over feedback-only designs in the frequency range from the corner frequency of S to the flexure resonance frequency. This perspective is further justified from the fact that the 2DOF mixed sensitivity optimal control designs (where the optimization problem formulation does not discriminate between feedforward and feedback components with respect to performance objectives) do not provide characteristically better performances than the 2DOF optimal robust model matching designs (where the optimization problem formulation places a robustness requirement on the feedback component and a bandwidth enhancement objective on the feedforward component).

5.3. Breaking barriers of feedback-only design

The 2DOF design is not bound by some fundamental limitations that constrain the feedback-only designs. For instance, the results in section 4 show that, in the 2DOF design, the tracking bandwidth ω_{BW} of the closed-loop device can be made larger than the roll-off frequency ω_T which determines the resolution. The corner frequency ω_{BW} can *never* be made larger than ω_T in feedback-only design, which suffers from a stricter trade-off between the resolution and the bandwidth. In optimal prefilter model matching control design, the closed-loop system has a bandwidth ω_{BW} of 214.5 Hz and a roll-off frequency ω_T of 60.1 Hz while the $\mathcal{H}_\infty$ feedback-only controller has a bandwidth of 49.4 Hz and the same roll-off frequency. This gives a case where 2DOF design achieves specifications that are *impossible* to attain with feedback-only design (see page 39 in [17]). In contrast to the results obtained by 2DOF design, the simultaneous improvement in the bandwidth and robustness at fixed resolution is typically difficult to achieve in feedback-only design.

5.4. Control design extensions

The 2DOF designs presented here can be easily extended to achieve a larger class of design specifications than presented in this paper, with trivial modifications to the designs presented here. For instance, we have not exploited the frequency content of the reference signals in our designs. This can be easily incorporated by writing the reference signal as $r = H(j\omega)r'$, where r' is an arbitrary signal (as used in our analysis) and the transfer function $H(j\omega)$ reflects the frequency content of the signal. Alternatively we can incorporate the frequency content of the reference signal through the model transfer functions T_{ref} or through weight transfer functions.

In this paper, we have demonstrated experiments that achieve typical trade-offs between performance objectives. We can achieve extreme trade-offs by suitably designing the reference transfer functions T_{ref} or weight transfer functions. For instance, very high resolution can be achieved by trading off severely on the tracking bandwidth. In [31], this trade-off (albeit in feedback-only design) yielded subnanometer resolution for the nanopositioning system described here. In fact, in this design the scanning speed for repetitive scans was not sacrificed by making the closed-loop device behave as a small-bandwidth device around the scanning frequency.

5.5. Integration of device and control design

In this paper, we have presented control designs assuming that the scanner G has already been designed. The integration of control design and device design can lead to a further increase in the closed-loop device performance. For instance, in [32] lighter positioning systems with single moving mass for multiple directions albeit with high cross-coupling effects are presented. The lighter mass results in higher flexure resonance frequencies and the control design decouples the cross interactions which has resulted in an overall superior device. The fundamental limitations developed in this paper can be used in the analysis of device design that will enable better closed-loop devices. For instance, the device design

can be altered and resulting changes in non-minimum phase zeros (or their locations) be monitored that will result in better closed-loop devices.

6. Conclusions

This paper describes 2DOF control methods for nanopositioning systems. In nanopositioning systems, along with high bandwidth and high resolution, robustness assumes great significance. A robust linear controller is needed to tackle the nonlinearities associated with piezoactuation and the changing flexure dynamics without having to design specific nonlinear controllers. The design goals of robustness, bandwidth, and resolution can be quantified in a straightforward manner in the framework of modern robust control. 2DOF control design which deals with the reference and the measured output signal separately has better capability to satisfy robustness and performance objectives. Three types of 2DOF control design (optimal prefilter model matching design, 2DOF mixed sensitivity synthesis, and 2DOF optimal robust model matching design) are described. The control designs when implemented on an AFM have given substantial improvements in bandwidth (as high as 330%) for the same resolution and robustness. Other performance objectives can be improved by appropriately designing the weight function or target transfer function. The trade-off imposed by fundamental limitations in 1DOF can be relaxed through 2DOF control. The experiments demonstrated that the 2DOF design presented here achieved design specifications that were analytically (and therefore practically) impossible for feedback-only designs. Here the comparisons have been made to feedback-only designs which by themselves have obtained significant improvements over commercial devices [11]. This paper presents a detailed analysis of fundamental limitations on the nanopositioning systems as well as the fundamental trade-offs between the performance objectives.

Acknowledgment

This work was supported by NSF Grant No. ECS 0449310 CAR.

References

[1] Bhushan B 1999 *Handbook of Micro/Nano Tribology* 2nd edn (Boca Raton, FL: CRC Press)
[2] Vettiger P *et al* 2002 The millipede-nanotechnology entering data storage *IEEE Trans. Nanotechnol.* **1** 39–55
[3] Sebastian A, Pantazi A, Cherubini G, Eleftheriou E, Lantz M A and Pozidis H 2005 Nanopositioning for probe storage *Proc. American Control Conf. (Portland, OR, June)* pp 4181–6
[4] Xu J, Lynch M, Huff J L, Mosher C, Vengasandra S, Ding G and Henderson E 2004 Microfabricated quill-type surface patterning tools for the creation of biological micro/nano arrays *IEEE Trans. Nanotechnol.* **6** 1387–2176
[5] Croft D, Shedd G and Devasia S 2000 Creep, hysteresis and vibration compensation for piezoactuators: atomic force microscopy application *Proc. American Control Conf. (Chicago, IL, June)* pp 2123–8
[6] Hatch A G, Smith R C, De T and Salapaka M V 2006 Construction and experimental implementation of a model-based inverse filter to attenuate hysteresis in ferroelectric transducers *IEEE Trans. Control Syst. Technol.* **14** 1058–69
[7] Smith R C, Hatch A G, De T, Salapaka M V, del Rosario R C H and Raye J K 2006 Model development for atomic force microscope stage mechanisms *SIAM J. Appl. Math.* **66** 1998–2026
[8] Ge P and Jouaneh M 1996 Tracking control of a piezoceramic actuator *IEEE Trans. Control Syst. Technol.* **4** 209–16
[9] Bhikkaji B, Ratnam M, Fleming A J and Moheimani S O R 2007 High-performance control of piezoelectric tube scanners *IEEE Trans. Control Syst. Technol.* **5** 853–66
[10] Daniele A, Salapaka M V, Salapaka S and Dahleh M 1999 Piezoelectric scanners for atomic force microscopes: design of lateral sensors, identification and control *Proc. American Control Conf. (San Diego, CA, June)* pp 253–7
[11] Salapaka S, Sebastian A, Cleveland J P and Salapaka M V 2002 High bandwidth nano-positioner: a robust control approach *Rev. Sci. Instrum.* **73** 3232–41
[12] Leang K and Devasia S 2002 Hysteresis, creep and vibration compensation for piezoactuators: feedback and feedforward control *Proc. 2nd IFAC Conf. on Mechatronic Systems* pp 283–9
[13] Schitter G, Allgower F and Stemmer A 2004 A new control strategy for high speed atomic force microscopy *Nanotechnology* **15** 108–14
[14] Zou Q and Devasia S 2004 Preview-based optimal inversion for output tracking: application to scanning tunneling microscopy *IEEE Trans. Control Syst. Technol.* **12** 375–86
[15] Aphale S S, Devasia S and Moheimani S O R 2008 High-bandwidth control of a piezoelectric nanopositioning stage in the presence of plant uncertainties *Nanotechnology* **19** 125503
[16] Helfrich B E, Lee C, Bristow D A, Xiaohui X, Dong J, Alleyne A G and Salapaka S 2008 Combined $\mathcal{H}_\infty$-feedback and iterative learning control design with application to nanopositioning systems *Proc. American Control Conf. (Seattle, WA, June)* pp 3893–900
[17] Skogestad S and Postlethwaite I 2005 *Multivariable Feedback Control, Analysis and Design* 2nd edn (New York: Wiley)
[18] Franklin G F, Powell J D and Workman M 1998 *Digital Control of Dynamic Systems* 3rd edn (Menlo Park, CA: Addison Wesley–Longman)
[19] Bode H 1945 *Network Analysis and Feedback Amplifier Design* (New York: Van Nostrand Reinhold)
[20] Freudenberg J S and Looze D P 1985 Right half-plane poles and zeros and design tradeoffs in feedback systems *IEEE Trans. Autom. Control* **30** 555–65
[21] Mohtadi C 1990 Bode's integral theorem for discrete-time systems *IEE Proc.* D **137** 57–66
[22] Doyle J C, Francis B A and Tannenbaum A R 1992 *Feedback Control Theory* (New York: Macmillan)
[23] Sebastian A and Salapaka S 2005 Design methodologies for robust nano-positioning *IEEE Trans. Control Syst. Technol.* **13** 868–76
[24] Glover K and McFarlane D 1989 Robust stabilization of normalized coprime factor plant descriptions with $\mathcal{H}_\infty$-bounded uncertainty *IEEE Trans. Autom. Control* **34** 821–30
[25] McFarlane D and Glover K 1992 A loop shaping design procedure using $\mathcal{H}_\infty$ synthesis *IEEE Trans. Autom. Control* **37** 759–69
[26] Hoyle D J, Hyde R A and Limebeer D J N 1991 An $\mathcal{H}_\infty$ approach to two degree of freedom design *Proc. IEEE Conf. on Decision and Control (Dec.)* pp 1581–5
[27] Limebeer D, Kasenally E and Perkins J 1993 On the design of robust two degree of freedom controllers *Automatica* **29** 157–68

[28] Vidyasagar M 1985 *Control System Synthesis: a Factorization Approach* (Cambridge, MA: MIT Press)
[29] Dullerud G E and Paganini F 2000 *A Course in Robust Control Theory: A Convex Approach* (*Texts in Applied Mathematics* vol 36) (Berlin: Springer)
[30] Lee C and Salapaka S 2008 Optimal model matching design for high bandwidth, high resolution positioning in AFM *Proc. 17th IFAC Conf. (Seoul, July)* pp 9230–5
[31] Shegaonkar A and Salapaka S 2007 Making high resolution positioning independent of scan rates: A feedback approach *Appl. Phys. Lett.* **91** 203513
[32] Dong J, Salapaka S and Ferreira P 2007 Robust mimo control of a parallel kinematics nano-positioner for high resolution high bandwidth tracking and repetitive tasks *Proc. Control and Decision Conf. (New Orleans, LA, Dec.)* pp 4495–500