Fast Robust Nanopositioning—A Linear-Matrix-Inequalities-Based Optimal Control Approach

Chibum Lee, *Student Member, IEEE*, and Srinivasa M. Salapaka, *Member, IEEE*

***Abstract*—This paper proposes a 2-DOF robust optimal control design method for achieving multiple objectives of resolution, bandwidth, and robustness to modeling uncertainties in nanopositioning systems. The main theoretical contribution of this paper is the formulation of a multiobjective 2-DOF optimal control problem in terms of linear matrix inequalities, which are then solved using standard convex optimization tools. The main distinguishing feature of this approach is the flexibility this method provides in formulating and solving the optimization problems that results in achieving a larger set of performance specifications. It facilitates solving of a certain class of mixed-norm optimization and pole-placement problems that arise naturally in nanopositioning systems. This methodology is demonstrated through experiments on a nanopositioning system that archives performance specifications, which are impossible with 1-DOF designs. Experimental results also demonstrate over 200% improvement in bandwidth of the resulting nanopositioning system over the optimal 1-DOF control designed for the same resolution and robustness specifications.**

***Index Terms*—High bandwidth, high resolution, linear matrix inequalities (LMIs), robust nanopositioning, 2-DOF control design.**

I. Introduction

NANOPOSITIONING is one of the critical requirements for nanoscience and technology. The interrogation and manipulation at the nanometer scale necessitate positioning systems to have equal or higher resolution. There is an added impetus on design of nanopositioning systems since they form the bottleneck in terms of speed and accuracy of most devices for nanoinvestigation, such as scanning probe microscopes. For example, the cantilever probe in an atomic force microscope (AFM) that measures atomic scale interaction between itself and the sample matter has natural frequencies in the order of 100 kHz, while typical positioning systems have bandwidths in 0.01–1 kHz range. Along with high precision, most nanoscientific studies and applications impose severe demands on tracking bandwidth and reliability in terms of repeatability of experiments. High tracking bandwidth is necessitated by applications that track dynamic evolution of nanoscale features, such as tracking cell-wall dynamics in cellular biology [1]; or high throughput studies, such as ultrahigh density media-storage applications [2], that require scanning areas with characteristic dimensions in micrometers and millimeters with nanoscale precision. The main causes for inconsistencies or lack of repeatability in experiments stem from modeling uncertainties, where complex effects of the environment on sensitive probes are not accurately accounted or typically ignored. The demand on reliability of positioning systems translates to specification of robustness to modeling uncertainties and operating conditions.

Typically, nanopositioning systems comprise a flexure stage that provides a frictionless motion through elastic deformation and piezoactuators that drive it. Piezoactuators are popular since they have no backlash, no wear and tear, have fast response, and are invulnerable to temperature changes and electric fields. However, they cause the nonlinear effects such as hysteresis and creep [3].

Lack of accurate models for nonlinear effects of piezoactuation and other uncertainties, such as thermal and humidity effects, along with severe demands on resolution and tracking bandwidth for nanopositioning systems has motivated substantial research in this area. These efforts can largely be characterized as those that aim at redesigning the flexure stage mechanisms and those that propose new control methods for existing stages. The most common redesign approach is to design smaller and stiffer flexure stages which result in higher resonant frequencies, and therefore provide better tracking bandwidths [4], [5]. However, the downside of this approach is that the resulting smaller stiffer devices have reduced traversal ranges. Many approaches have been proposed that study and develop models for nonlinear effects such as hysteresis, and implement feedforward control schemes such as iterative learning control (ILC) to compensate for them [3], [6], [7]. The use of charge instead of voltage in driving piezoelectric tubes is another way for reducing hysteresis [8]. Feedback control methods, on the other hand, study the dominant frequency ranges of the uncertainties in the model and implement feedback laws that make the device insensitive to them while guaranteeing a specified performance [9]–[12]. For instance, in [10], optimal $\mathcal{H}_\infty$ control problems that reflect resolution, bandwidth, and robustness are formulated and solved to achieve large improvements in tracking performance and robustness to operating conditions. A concerted effort of device and control design is presented in [13], where a lighter (and therefore, faster) device is designed by having a single moving mass and the responsibility of compensating mechanical coupling is delegated to the multi-input–multi-output (MIMO) control design.

Manuscript received October 22, 2008; revised March 21, 2009. Current version published August 14, 2009. Recommended by Guest Editor A. Ferreira. This work was supported by National Science Foundation (NSF) under Grant ECS 0449310 CAR and Grant CMMI 08-00863.

The authors are with the Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: salapaka@illinois.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TMECH.2009.2023903

More recently, 2-DOF designs that combine the feedforward and feedback strategies have been reported [14]–[17]. 2-DOF design schemes are discussed in [14] where optimal inversion control (as well as ILC schemes [18]) are used along with the feedback control to make improvements on the feedback-only design. A polynomial-based feedback controller is designed in [15] to account for pole-location uncertainty along with inversion-based feedforward controller, which provides improved tracking for raster-scan applications. The fundamental limitations on achievable performance specifications for 1-DOF systems are discussed in [16] and [17], and compared to the 2-DOF systems. A 2-DOF optimal control problem based on $\mathcal{H}_\infty$ stacked sensitivity framework [19] is formulated, solved, and implemented. Relative roles of feedforward and feedback components are analyzed, and, analytical as well as experimental comparisons are made between the 1-DOF and 2-DOF systems. In the control systems literature, the design of 2-DOF systems has been topic of much research, which has been resolved using various techniques, each with their unique advantages. For instance, in [20], 2-DOF control design is parameterized in terms of two independent stable rational parameters (each related to feedforward and feedback) and analyzed with this parameterization. The 2-DOF control design, with extra emphasis on robustness, is developed in [21] based on the integration of the Glover and McFarlane loopshaping and model-matching that makes the closed-loop responses close to a specified response.

In this paper, we propose a new optimal control method for designing 2-DOF feedback laws. This paper shares the same goals as in [16], [17], [22], and [23], in terms of achieving multiple objectives of resolution, bandwidth, and robustness to modeling uncertainties in nanopositioning systems. The main theoretical contribution of this paper is the formulation of this multiobjective 2-DOF optimal control problem in terms of linear matrix inequalities (LMIs), which are then solved using standard convex optimization tools. The 2-DOF optimal control problem is cast in such a way that it can be adapted into the LMI-based framework proposed in [24]. In this sense, the paper results in a solution for 2-DOF system design problem that is not restricted to nanopositioning applications. However, the relevance of this methodology to nanopositioning systems is twofold. Solving the same optimization problems in the stacked $\mathcal{H}_\infty$ sensitivity framework presented in [17] results in more complex controllers, i.e., with higher orders, than the method presented in this paper. Since the sampling frequency, and hence the controller bandwidth, is limited by the complexity of the controller, the proposed method gives more practical solutions. The main advantage, though, is the flexibility this method provides in formulating the optimization problems. It facilitates solving of a class of mixed-norm optimization and pole-placement problems that arise naturally in nanopositioning systems. Therefore, it results in obtaining a larger class of performance specifications than those obtained in [17]. It should be emphasized that these advantages distinguish the proposed methodology over the designs presented in [16] and [17]. Since they solve similar optimization problems, the proposed method inherits all the advantages of the designs presented in [16] and [17]. For instance, the experimental results demonstrate that the resulting controllers achieve much larger bandwidths than the feedforward-only and feedback-only designs (over 200% larger bandwidth for the same resolution and robustness measures), and in fact, they achieve specifications that are analytically impossible with 1-DOF linear controllers.

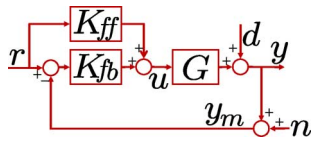


Fig. 1. 2-DOF control architectures. The feedforward–feedback scheme where the actuation signal $u = K_{ff}r + K_{fb}(r - y_m)$.

This paper is organized as follows. Section II provides the control design objectives and formulation of a related 2-DOF optimal control problem. In Section III, this optimal control problem is formulated in terms of LMIs. In Section IV, the resulting 2-DOF design is demonstrated on a scanner of an AFM and substantiated through experimental results. An analysis and discussion of the proposed methodology and its implementation are presented in Section V. We conclude with final observations in Section VI.

II. 2-DOF Control Framework and Objectives

Among various architectures for 2-DOF control, we consider the scheme shown in Fig. 1. In this figure, G is the transfer function of the *scanner* comprising of the actuator, the flexure stage, and the sensor. The signal y represents its output, the flexure stage displacement (scaled by a sensor constant), and the signal u represents its input given to the actuator. The signal r represents the command signal that the positioning system needs to track, d represents the *mechanical noise*—the effects of dynamics that are not incorporated in the model G, n represents the sensor noise, $y_m = y + n$ represents the noisy measurement, and K_{ff} and K_{fb}, respectively, represent the feedforward and feedback control transfer functions. The main objective for the design of the controllers K_{ff} and K_{fb} is to make the *tracking error* small for the largest possible range of frequencies despite the uncertainties. In other words, the performance of a nanopositioning system is characterized by its positioning resolution, tracking bandwidth, and robustness to modeling uncertainties.

For a given 2-DOF controller with K_{ff} and K_{fb}, the closed-loop signals are given by

$$\begin{aligned} y &= SG(K_{ff} + K_{fb})r - Tn + Sd \\ e &= S(1 - GK_{ff})r + Tn - Sd \\ u &= S(K_{ff} + K_{fb})r - SK_{fb}n - SK_{fb}d \end{aligned} \tag{1}$$

where $S{=}(1 + GK_{fb})^{-1}$, $T{=}1 - S{=}(1 + GK_{fb})^{-1}GK_{fb}$ and e represents tracking error. In the feedback-only design, the case where $K_{ff} = 0$, the transfer function from n to e and from r to e cannot be simultaneously made small since $S + T = 1$; whereas in 2-DOF control design, this limitation is relaxed due to the extra freedom from K_{ff}. We use T_{yr} and S_{er} to denote the transfer functions from r to y and from r to e, respectively, i.e., $S_{er} = S(1 - GK_{ff})$, $T_{yr} = SG(K_{ff} + K_{fb})$.

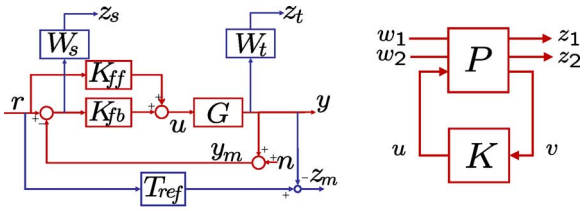


Fig. 2. (a) 2-DOF control design configuration. The signals z_m, z_s, and z_t represent the deviation from target response, the weighted tracking error with noise, and the weighted system output. The weights W_s and W_t are chosen to reflect the design specifications of robust stability, disturbance rejection, and noise attenuation. The target response function $T_{\rm ref}$ is chosen to specify the performance of closed-loop response to reference signal. In this 2-DOF framework, the controller has access to both the reference r and the measured error $r-n-y$. (b) General framework for multiobjective optimization problem: 2-DOF multiobjective control problem in (3) is to design a controller that minimizes the sum of ρ times the transfer function from $w_1=r$ to $z_1=z_m$ and the transfer function from $w_2=n$ to $z_2=[z_s\ z_t]^T$.

The performance objectives are characterized in terms of error e in (1). Small tracking error can be achieved by designing S_{er}, T, and S small in those frequency ranges, where the frequency contents of r, n, and d, respectively, are dominant. Broadband tracking is measured by how small S_{er} is at operating frequency and its bandwidth $\omega_{\rm BW}$, which is characterized by the frequency at which the magnitude $|S_{er}|$ crosses $1/\sqrt{2}$ [25]. The resolution of the closed-loop system is determined by the effect of noise from Tn in (1). Therefore, the small roll-off frequency ω_T at which the magnitude $|T|$ crosses $1/\sqrt{2}$ and high roll-off rates for T guarantee good resolution. Robustness to modeling uncertainties can be measured by $\|S\|_\infty$, which can be motivated by the effect of mechanical noise d. Low (near 1) values of $\|S\|_\infty$ guarantee good gain and phase margin since $GM \geq \|S\|_\infty/(\|S\|_\infty-1)$ and $PM \geq 2\sin^{-1}1/\|S\|_\infty$ [25]. Note that the transfer functions S and T characterize the robustness and resolution objectives in *both* 1-DOF and 2-DOF designs. Therefore, we cannot expect advantage of noise or disturbance rejection from 2-DOF design over 1-DOF design. The main advantage of 2-DOF over 1-DOF design is in extending the bandwidth of the nanopositioning system.

III. 2-DOF Control Design

In order to reflect the performance objectives discussed in previous section into optimal control setting, the configuration in Fig. 2(a) is considered. The main idea in this setup is to shape the closed-loop transfer functions S and T with weighting functions W_s and W_t to achieve robust stability, disturbance rejection, and noise attenuation, and to make the close-loop response close to a target reference response $T_{\rm ref}r$. The regulated outputs are chosen as $z_m = T_{\rm ref}r - y$, the deviation from target response, $z_s = W_s(r-n-y)$, the weighted tracking error with noise, and $z_t = W_t y$, the weighted system output.

Using (1), the closed-loop matrix transfer function from the exogenous variables $w=[r\ n]^T$ to the regulated variables $z=[z_m\ z_s\ z_t]^T$ is given by

$$\begin{bmatrix} z_m \\ z_s \\ z_t \end{bmatrix} = \underbrace{\begin{bmatrix} T_{\rm ref}-T_{yr} & T \\ W_sS_{er} & -W_sS \\ W_tT_{yr} & -W_tT \end{bmatrix}}_{=\Phi} \begin{bmatrix} r \\ n \end{bmatrix}. \tag{2}$$

Note that the transfer functions W_sS_{er}, W_sS, W_sT_{yr}, and W_tT, which are elements of matrix Φ in (2), cannot be simultaneously minimized due to algebraic limitations of $S+T=1$ and $S_{er}+T_{yr}=1$. These constraints severely restrict the feasible space of controllers that make $\|\Phi\|$ small. In our approach, we impose specifications directly on certain transfer functions instead of posing the problem in terms of regulated variables. More specifically, we target the transfer functions $T_{\rm ref}-T_{yr}$, W_sS, and W_tT. The cost objectives to be minimized are $T_{\rm ref}-T_{yr}$, which is the transfer function from r to z_m and $-[W_sS\ W_tT]^T$, which is the transfer function from n to $[z_s\ z_t]^T$. The optimization problem is stated as

$$\min_{K\in\mathcal{K}} \rho\,\|T_{\rm ref}-T_{yr}\|_\infty + \left\| \begin{matrix} W_sS \\ W_tT \end{matrix} \right\|_\infty \tag{3}$$

where $K=[K_{ff}\ K_{fb}]$, $\mathcal{K}$ is the set of stabilizing controllers, and the parameter ρ reflects the user-defined relative weight between model-matching and robust-performance objectives. Fig. 2(b) shows the general framework of multiobjective optimization in (3) corresponding to Fig. 2(a). The determination of the controller K that solves this problem is summarized in Theorem 1. The reinterpretation of optimization problem (3) in terms of Fig. 2(b) helps in casting it in an LMI-based framework proposed in [24].

Theorem 1: For the 2-DOF system given in Fig. 2(a) the following hold.

a) The multiobjective optimization problem in (3) has a solution when the following convex minimization problem has a solution:

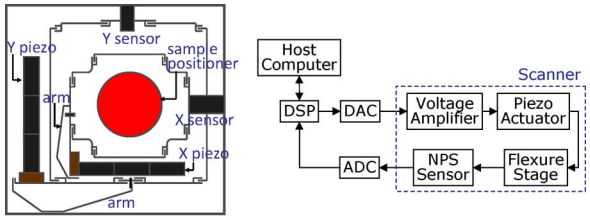
$$\begin{aligned} &\text{minimize } c(X) \\ &\text{subject to } F_1(X)<0, F_2(X)<0, \text{ and } F_3(X)<0 \end{aligned} \tag{4}$$

where $X=(\hat{A},\hat{B},\hat{C},\hat{D},R,S,\gamma_1,\gamma_2)$, $c(X)=\rho\gamma_1+\gamma_2$, and $F_1(X)$, $F_2(X)$, $F_3(X)$ are as shown at the bottom of the next page.

b) There exists a transformation $\Gamma:\mathcal{X}\to\mathcal{K}$ such that $K=\Gamma(X)$, where X is the solution of (4), where $\mathcal{K}$ is space of stabilizing control transfer functions.

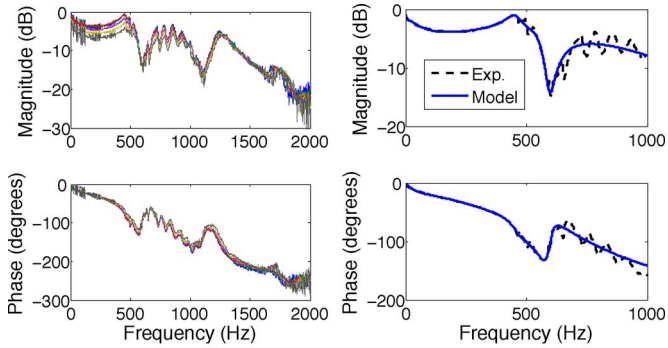
Here, $F_j(X)$ are symmetric matrices for $j=\{1,2,3\}$, the operation $\mathcal{Q}(Z)$ denotes $Z+Z^T$, matrices A,B,C, and D form state-space realization of the system transfer function P, as shown in Fig. 2, the matrices $\hat{A},\hat{B},\hat{C}$, and $\hat{D}$ determine the state-space realizations of controller K, and matrices B_j,C_j,D_j,E_j, and H_j for $j\in\{1,2\}$, and R and S are completely determined by A,B,C, and D. The set $\mathcal{X}=\{X,\ \text{s.t. } X=(\hat{A},\hat{B},\hat{C},\hat{D},R,S,\gamma_1,\gamma_2)\}$ is the feasible domain space for the variable X defined in (4). (The details of definitions of all these parameters are given in the proof.)

Proof: See the Appendix. ■



(a) (b)

Fig. 3. (a) Schematic of flexure scanner. The sample is placed on the central block of the flexure stage that is driven by the x-piezo, which, in turn, is driven by the y-piezo. The x- and y sensors measure the stage position. (b) Block diagram schematic of a closed-loop nanopositioning system. The scanner consists of a input signal amplifier, a piezoactuator, a flexure stage, and a sensor. The control law is implemented using a host computer and a DSP board where sensor readings are discretized and processed, and the actuation signals are converted to analog signals and applied to the scanner system.



(a) (b)

Fig. 4. (a) Experimental frequency responses at various dc offsets. (b) Nominal frequency response (dashed) and model frequency response (solid).

IV. Demonstration of Control Design

A. Device Description

The nanopositioning system that we use in this paper is a 2-D flexure scanner of MFP-3D developed by Asylum Research, Inc., Santa Barbara, CA. The flexure scanner consists of two stages (fast axis "X" stage is located on top of the slow axis "Y" stage) that provide motion through their elastic deformation, stack-piezos for actuation, and linear variable differential transformer (LVDT) based nanopositioning-system (NPS) sensors [see Fig. 3(a)]. These actuators lead to a travel range of $90\ \mu$m in closed-loop in both directions. The NPS sensors have noise less than 0.6 nm (deviation) over 0.1–1 kHz bandwidth.

A block diagram schematic of the closed-loop scanner system is shown in Fig. 3(b). The control law is discretized and implemented on a Texas instrument TMS320C6713 DSP with 16-bit A/D and 16-bit D/A channels. Physical modeling of the device is difficult due to its complicated structural design and poorly understood piezoactuation phenomena, and any attempt in that direction results in significant modeling uncertainty. Therefore, identification techniques were used to derive linear models about an operating point (dc offset), where the sensor output gives a "zero" reading corresponding to "zero" input to the piezoactuators.

The scanner is identified from a sine-sweep with 10 mV constant amplitude for frequency range of 1 Hz to 2 kHz by using HP 35670A dynamic signal analyzer. The small amplitude of voltage (10 mV over -10 to 10 V input range) is applied to the input of each axis to obtain the linear model at the operating point. From the identification results, X and Y crosstalk are seen to be relatively small ($\max|G_{xy}| = -17.76$ dB) since, by design, X and Y stages 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 slow. Hence, there is a greater emphasis on the control designs for the X stage, which is presented in this paper.

The identification is repeated to obtain frequency responses of the system by giving various dc offsets spanning the range of operation of the device [see Fig. 4(a)]. The variation in these responses is indicative of the modeling errors arising from the nonlinear effects of piezoactuation along with other sources of uncertainties. In addition, it is 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 operating point, which is at dc offset corresponding to 0 V output value. Fig. 4(b) shows the Bode diagram of fitted mathematical model with

$$F_1(X) = \begin{bmatrix} \mathcal{Q}(AR + B\hat{C}) & (*) & (*) & (*) \\ \hat{A} + (A + B\hat{D}C)^T & \mathcal{Q}(SA + \hat{B}C) & (*) & (*) \\ (B_1 + B\hat{D}H_1)^T & B_1^T S + H_1^T \hat{B}^T & -\gamma_1 I & (*) \\ C_1 R + E_1 \hat{C} & C_1 + E_1 \hat{D} C & D_1 & -\gamma_1 I \end{bmatrix} < 0$$

$$F_2(X) = \begin{bmatrix} \mathcal{Q}(AR + B\hat{C}) & (*) & (*) & (*) \\ \hat{A} + (A + B\hat{D}C)^T & \mathcal{Q}(SA + \hat{B}C) & (*) & (*) \\ (B_2 + B\hat{D}H_2)^T & B_2^T S + H_2^T \hat{B}^T & -\gamma_2 I & (*) \\ C_2 R + E_2 \hat{C} & C_2 + E_2 \hat{D} C & D_2 & -\gamma_2 I \end{bmatrix} < 0$$

$$F_3(X) = \begin{bmatrix} R & (*) \\ 1 & S \end{bmatrix} > 0$$

nominal experimental result. Weighted iterative least square fitting is performed over 0–1 kHz and the reduction through balanced realization [26] resulted in the following seventh-order model:

$$G_{xx}(s) = \frac{-1122.3157(s-1.152\times10^4)(s+543)}{(s+390.3)(s^2+470.9s+8.352\times10^6)} \times \frac{(s^2+587.2s+8.628\times10^6)(s^2+226.5s+1.407\times10^7)}{(s^2+689.4s+1.315\times10^7)(s^2+4950s+2.44\times10^7)}. \quad (5)$$

This seventh-order model did not capture dynamics beyond 500 Hz, as shown in Fig. 4(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 that cannot be accommodated by the processor with short sampling time. This modeling uncertainty from using low-order model is accounted by imposing the requirement of robustness on the control design.

B. Experimental Results

The controller resulting from the design described in Section III is applied to G_{xx} with $T_{\rm ref} = 1/(0.0003s+1)$, $W_s = (0.5s+394.8)/(s+0.3948)$, and $W_t = (100s+9.475\times10^4)/(s+1.184\times10^5)$, which reflect the performance objectives of high bandwidth, high resolution, and robustness to modeling errors. High-bandwidth objective requires that $|S_{er}|$ is small for wide range of operating frequencies. It is imposed by matching T_{yr} $(=1-S_{er})$ with $T_{\rm ref}$. High-resolution objective requires the roll-off frequency ω_T to be small, which is reflected in W_t. The robustness requirement is reflected in W_s, which enforces $\|S\|_\infty \le 2$. The weight of $\rho = 20$ is chosen for the multiobjective minimization.

The resulting 2-DOF controller from this design is given by

$$K_{ff} = \frac{2.746\times10^{14}(s+1.184\times10^5)(s+0.7883)(s+390.3)}{(s+3.131\times10^{11})(s+1.296\times10^8)(s+7817)(s+3333)} \times \frac{(s^2+470.9s+8.352\times10^6)(s^2+689.4s+1.315\times10^7)}{(s+543)(s+0.789)(s^2+587.2s+8.628\times10^6)} \times \frac{(s^2+4950s+2.44\times10^7)}{(s^2+226.5s+1.407\times10^6)}$$

$$K_{fb} = \frac{6.9497\times10^{13}(s+1.184\times10^5)(s+3333)(s+390.3)}{(s+3.131\times10^{11})(s+1.296\times10^8)(s+7817)(s+3333)} \times \frac{(s^2+470.9s+8.352\times10^6)(s^2+689.4s+1.315\times10^7)}{(s+543)(s+0.789)(s^2+587.2s+8.628\times10^6)} \times \frac{(s^2+4950s+2.44\times10^7)}{(s^2+226.5s+1.407\times10^6)}. \quad (6)$$

The corresponding closed-loop transfer functions are shown in Fig. 5.

The 2-DOF control laws obtained from multiobjective synthesis were implemented. The experimental results are shown in Fig. 6. The experimentally obtained tracking bandwidth $\omega_{\rm BW}$ is 161 Hz [from $\|S_{er}\|$ in Fig. 6(c)], the robustness measure $\|S\|_\infty$ is 1.06 [as shown in Fig. 6(a)], and roll-off frequency ω_T, which determines resolution, is 57.5 Hz [from Fig. 6(b)].

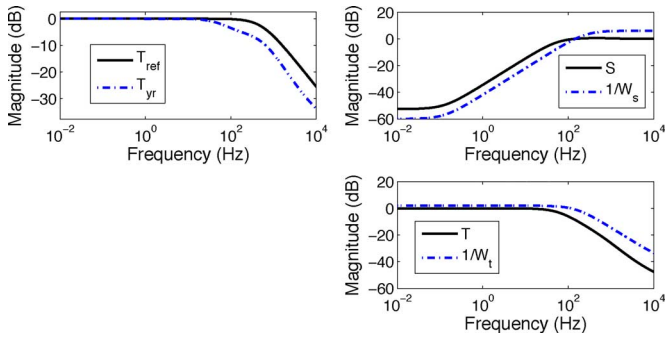


Fig. 5. Multiobjective synthesis of 2-DOF control using LMI. The minimization results corresponding to weight $\rho = 20$ are $\|T_{\rm ref} - T_{yr}\|_\infty = 0.526$ and $\|[W_s S\ W_t T]^T\|_\infty = 2.48$.

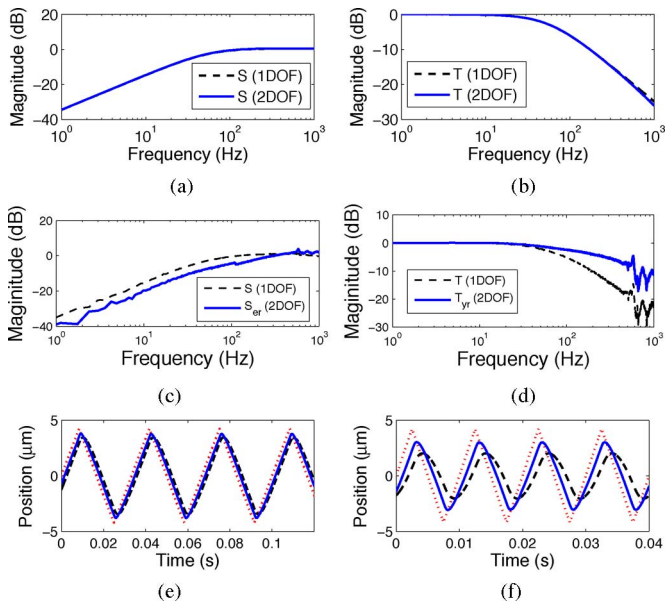


Fig. 6. Comparison of 1-DOF (feedback-only) $\mathcal{H}_\infty$ stacked-synthesis and 2-DOF multiobjective synthesis control design. (a) and (b) Comparison of $|S|$ and $|T|$ in 1-DOF (dashed lines) and 2-DOF (solid lines) control designs. (c) and (d) Comparison of $|S|$ and $|T|$ in 1-DOF (dashed lines), and $|S_{er}|$ and $|T_{yr}|$ from 2-DOF (solid lines) control designs. The 2-DOF multiobjective synthesis control achieves about 216% improvement in the tracking bandwidth. (e) and (f) Triangular reference signals (dotted lines) with frequencies 30 and 100 Hz are shown. The tracking performance of 2-DOF design (solid lines) is better than the 1-DOF design (dashed lines), especially at high frequencies.

1) Comparison with feedback-only design: For comparison with *optimal* feedback-only 1-DOF design, the feedback-only controller was designed by S/T stacked sensitivity synthesis such that the closed-loop system has transfer functions T and S similar to those in Fig. 6, which implies resolution and robustness measures are identical to the 2-DOF design.

The result of $\mathcal{H}_\infty$ synthesis yielded the following ninth-order control law:

$$K_{1D} = \frac{6.0151\times 10^{13}(s+2.531\times 10^5)(s+390.3)}{(s+1.471\times 10^{13})(s+3.292\times 10^6)(s+1.203\times 10^4)} \times \frac{(s^2+470.9s+8.352\times 10^6)(s^2+689.4s+1.315\times 10^7)}{(s+543)(s+0.8733)(s^2+587.2s+8.628\times 10^6)} \times \frac{(s^2+4950s+2.44\times 10^7)}{(s^2+226.6s+1.407\times 10^6)}. \quad (7)$$

The experiment results in $\|S\|_\infty = 1.17$, $\omega_T = 63.7$ Hz, and $\omega_{\mathrm{BW}} = 51.0$ Hz. There is an improvement of 216% in bandwidth, for nearly the same values of resolution and robustness, when compared to the optimal feedback-only design. To demonstrate the tracking performance, triangular signals at different frequencies were given as reference signals and the corresponding outputs from 2-DOF and 1-DOF designs are plotted in Fig. 6(e) and (f). The tracking error from 2-DOF design is 43% less than the error in the 1-DOF design for the 100 Hz reference signal. Both the 1-DOF and 2-DOF designs are solved for the same specifications on resolution and robustness (for input-to-error gain 1% at low frequencies, roll-off frequency ω_T for $T \approx 60$ Hz, and $\|S\|_\infty < 2$). Even better improvements in bandwidth can be achieved by relaxing specification on robustness or resolution in the optimal control problem (3).

Remark: Note that the 1-DOF control that we are comparing with is *optimal* in the sense that it hits the boundaries of achievable performance objectives of resolution, bandwidth, and robustness, i.e., *any* change in the specifications for better resolution and robustness in the 1-DOF optimization, leads to reduction in bandwidth. It should also be remarked that this 1-DOF design, by itself, is a significant improvement over other 1-DOF designs (in [10], we show over 60 times improvement in bandwidth over existing commercial designs). Therefore, showing an improvement of over 216% in bandwidth for same resolution and robustness specifications is a significant improvement over the *best* 1-DOF design.

2) Practical elimination of hysteresis through proposed design: Fig. 7(a) and (b) demonstrates practical elimination of hysteresis and justification of the linear model. As clearly seen from these experimental results, the closed-loop system gives desired (linear) motion for the entire range and it is insensitive to the error from model reduction assumed in the open-loop model.

V. Analysis and Discussion

A. Flexibility of Design Objective—Advantages From LMIs

One of the principal advantages of the LMI-based approach is that it allows *simultaneous* incorporation of different metrics (such as $\mathcal{H}_\infty$, $\mathcal{H}_2$) on performance constraints as well as including design requirements such as specifications on passivity, asymptotic tracking, and pole locations [24], [27]. This generality of approach, especially in incorporation of mixed-norm optimization problems, proves very useful in the context of nanopositioning. The measure on robustness is well characterized by infinity norm, but the natural metrics on resolution and bandwidth can vary depending on the applications and specific requirements. For instance, the metric on model mismatch can be given in terms of $\mathcal{H}_\infty$ norm as we have considered in this paper or $\mathcal{H}_2$ norm depending on whether we are interested in the "worst-case" or average performance. Thus, various specifications impose different optimization problems, say in the form

$$\min_{K\in\mathcal{K}} \rho\|T_{\mathrm{ref}} - T_{yr}\|_\alpha + \|T\|_\beta \ \text{ subject to } \|W_s S\|_\gamma < 1$$

where α, β, and γ can be two or ∞ norms.

This approach also allows for interchanging cost and constraints in an optimization problem. For instance, one can solve a problem where the cost function is maximization of robustness for given specifications of model mismatch and noise attenuation, i.e., one can solve for problems of the form

$$\min_{K\in\mathcal{K}} \|S\|_\infty \text{ subject to } \|T_{\mathrm{ref}} - T_{yr}\|_\infty < m_1 \text{ and } \|T\|_\infty < m_2.$$

Thus, this approach makes possible addressing a variety of nanopositioning applications with diverse objectives.

It should be emphasized that this approach imposes a technical condition (17) to solve the optimization problem. However, this condition adds conservatism to the solution in the sense that it minimizes over a subset of $\mathcal{K}$ instead of the entire $\mathcal{K}$. An area of active research is on relaxing this technical condition [28], [29].

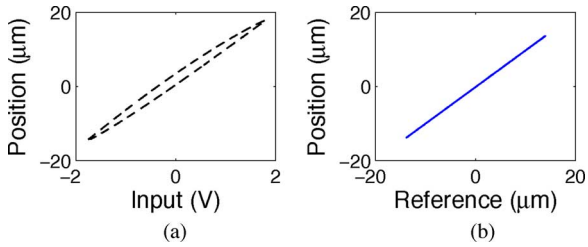


Fig. 7. Comparison of hysteresis in open- and closed-loop configuration. (a) Open-loop hysteresis: the maximum output hysteresis of 3.06 μm (9.4%) and the maximum input hysteresis of 0.32 V (9.0%) for actuation voltage from -1.8 to 1.8 V. (b) Closed-loop hysteresis: the maximum output and input hysteresis of 46 nm (0.2%) for reference command of 14 μm trace and retrace.

B. Comparison With 2-DOF Stacked Sensitivity Synthesis

Most of the existing 2-DOF designs cited in Section I are motivated from performance objectives that require *a priori* delegation of responsibilities between feedforward and feedback components, where they are sequentially designed. However, in our earlier research, 2-DOF stacked sensitivity synthesis was proposed for designing simultaneously feedback and feedforward controllers with configuration of [17, Fig. 8].

The corresponding closed-loop matrix transfer function for the optimization problem is given as

$$\begin{bmatrix} z_s \\ z_t \\ z_u \end{bmatrix} = \underbrace{\begin{bmatrix} W_s S_{er} W_r & -W_s S W_n \\ W_t T_{yr} W_r & -W_t T W_n \\ W_u S(K_{ff}+K_{fb}) W_r & -W_u K_{fb} S W_n \end{bmatrix}}_{=\Phi} \begin{bmatrix} r \\ n \end{bmatrix}. \quad (8)$$

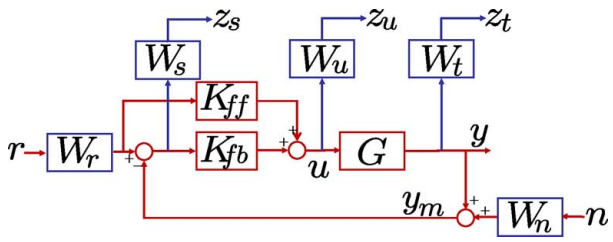


Fig. 8. 2-DOF control stacked-synthesis design configuration. All notations are similar with Fig. 1 except z_u, W_u, W_r, and W_n. The signal z_u represents the control signal and the weight W_u is chosen to reflect the design specifications of saturation limits on the control signal. Two weight functions W_r and W_n are introduced to carry the information content of input signals r and n, respectively.

In this formulation, the main challenge of design comes from upper two rows of Φ. This problem requires a design of complicated weight functions W_r, W_n, W_s, and W_t to shape the following closed transfer functions: S_{er} with W_sW_r, S with W_sW_n, T_{yr} with W_tW_r, and T with W_tW_n (the details can be found in [17]).

Since the resulting weight functions W_r and W_n are of high order, the controller obtained from the $\mathcal{H}_\infty$ results in a high-order controller. In contrast, multiobjective 2-DOF control design presented in this paper results in a lower order controller and does not require design of special weight functions.

Another important characteristic that distinguishes our LMI-based design from other existing 2-DOF designs is the added flexibility in specification of performance objectives, as described in Section V-A. However, ILC-based techniques can be incorporated in the proposed design to obtain further improvements (as discussed in [18]).

VI. Conclusion

This paper has described 2-DOF robust optimal control design and its implementation for nanopositioning systems, which simultaneously achieves high resolution, high bandwidth, and robustness to model uncertainties. To achieve these multiple objectives, a new 2-DOF problem was formulated as an LMI-based convex optimization problem. In this approach, a large class of performance objectives can be achieved by employing different norms on different objectives and specifying constraints that reflect design constraints such as pole locations, passivity, and other performance constraints. The controller is obtained by standard convex optimization tools. This methodology is demonstrated through experiments on a nanopositioning system. The main contribution of feedforward component is in the frequency range where S cannot be made small due to limitations such as the Bode integral law. The feedforward components provide performance *enhancements* from corner frequency of S to a little beyond flexure stage resonance frequency. The experiment results demonstrate 200% improvement in bandwidth over optimal 1-DOF designs and achieve specifications that are impossible in a 1-DOF framework.

Appendix

Lemma 1 (Kalman–Yakubovich–Popov): For

$$\Phi(s) = \left[\begin{array}{c|c} \bar{A} & \bar{B} \\ \hline \bar{C} & \bar{D} \end{array}\right]$$

the matrix $\bar{A}$ is Hurwitz and $\|\Phi\|_\infty < \gamma$ if and only if there exist a matrix $P > 0$ such that

$$\begin{bmatrix} \bar{A}^T P + P\bar{A} & (*) & (*) \\ \bar{B}^T P & -\gamma I & (*) \\ \bar{C} & \bar{D} & -\gamma I \end{bmatrix} < 0. \tag{9}$$

Proof: See [26] for the proof. ■

Proof of Theorem 1

Proof for (a): We recast the 2-DOF problem in a form that can use the framework developed in [24]. Note that the the generalized matrix P in Fig. 2(b) is given by

$$\begin{bmatrix} z_m \\ z_s \\ z_t \\ \hline r \\ r-n-y \end{bmatrix} = \underbrace{\left[\begin{array}{cc|c} T_{\text{ref}} & 0 & -G \\ W_s & -W_s & W_sG \\ 0 & 0 & W_tG \\ \hline I & 0 & 0 \\ I & -I & -G \end{array}\right]}_{=:P} \begin{bmatrix} r \\ n \\ u \end{bmatrix}. \tag{10}$$

We represent the state-space realization of P by

$$P = \left[\begin{array}{c|cc} A & B_w & B \\ \hline C_z & D_{zw} & D_z \\ C & D_w & 0 \end{array}\right] \tag{11}$$

and the controller by

$$K(s) = [K_{ff}(s)\ K_{fb}(s)] = \left[\begin{array}{c|c} A_k & B_k \\ \hline C_k & D_k \end{array}\right]. \tag{12}$$

In terms of these realizations, the overall closed-loop transfer Φ in (2) is then derived as

$$\begin{aligned} \Phi &= \left[\begin{array}{cc|c} A+BD_kC & BC_k & B_w+BD_kD_w \\ B_kC & A_k & B_kD_w \\ \hline C_z+D_zD_kC & D_zC_k & D_{zw}+D_zD_kDw \end{array}\right] \\ &=: \left[\begin{array}{c|c} \bar{A} & \bar{B} \\ \hline \bar{C} & \bar{D} \end{array}\right]. \end{aligned} \tag{13}$$

We represent the transfer functions from w_j to z_j in Fig. 2(b) by Φ_j's $(j = 1, 2)$, and the bounds on their $\mathcal{H}_\infty$ norms by γ_j. (The transfer function $\Phi_1 = T_{\text{ref}} - T_{yr}$ represents the model mismatch and $\Phi_2 = -[W_sS\ W_tT]^T$ reflects the robustness and resolution objectives.) Note that these are the transfer functions of interest that we seek to make small in our approach as evident by the cost function in (4). Each of these transfer functions can be written in terms of the matrix transfer function Φ in (2) as

$\Phi_j = L_j \Phi R_j$, where the matrices L_j and R_j are chosen as

$$L_1 = [\,1 \quad 0 \quad 0\,], \qquad R_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
$$L_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \text{and} \qquad R_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{14}$$

This decomposition of Φ provides a way for using results given in [24]. Let $B_j = B_w R_j$, $C_j = L_j C_z$, $D_j = L_j D_{zw} R_j$, $E_j = L_j D_z$, and $H_j = D_w R_j$ $(j = 1, 2)$ then

$$\begin{aligned} \Phi_j &= \left[\begin{array}{c|c} \bar{A} & \bar{B} R_j \\ \hline L_j \bar{C} & L_j \bar{D} R_j \end{array}\right] \\ &= \left[\begin{array}{cc|c} A + B D_k C & B C_k & B_j + B D_k H_j \\ B_k C & A_k & B_k H_j \\ \hline C_j + E_j D_k C & E_j C_k & D_j + E_j D_k H_j \end{array}\right] \\ &=: \left[\begin{array}{c|c} \bar{A} & \bar{B}_j \\ \hline \bar{C}_j & \bar{D}_j \end{array}\right], \qquad \text{for } j \in \{1, 2\}. \end{aligned} \tag{15}$$

The condition for $\mathcal{H}_\infty$ performance $\|\Phi_j\| < \gamma_j$ can be imposed by the Lemma 1. Hence, equivalent problem to this performance condition is determining P_j such that

$$\begin{gathered} P_j > 0 \\ \Theta_j := \begin{bmatrix} \bar{A}^T P_j + P_j \bar{A} & P_j \bar{B}_j & \bar{C}_j{}^T \\ \bar{B}_j{}^T P_j & -\gamma_j I & \bar{D}_j{}^T \\ \bar{C}_j & \bar{D}_j & -\gamma_j I \end{bmatrix} < 0. \end{gathered} \tag{16}$$

Note that the matrix inequality (16) is not linear in terms of the actual design variables (A_k, B_k, C_k, D_k). However, they can be converted into LMIs if we impose the condition

$$P_1 = P_2 = P \tag{17}$$

and change the variables through an appropriate congruence transform. Assumption (17) brings the conservatism in design but it recovers linearity of variables.

A short sketch of construction of the transform that makes (16) LMIs in design variables is given shortly. We decompose the unknown positive-definite matrix P and its inverse as

$$P = \begin{bmatrix} S & N \\ N^T & ? \end{bmatrix}, \qquad P^{-1} = \begin{bmatrix} R & M^T \\ M^T & ? \end{bmatrix} \tag{18}$$

where the terms shown as ? are unimportant. The transformation is defined in terms of the following matrices

$$\Pi_1 = \begin{bmatrix} R & I \\ M^T & 0 \end{bmatrix}, \qquad \Pi_2 = \begin{bmatrix} I & S \\ 0 & N^T \end{bmatrix}. \tag{19}$$

Note that $PP^{-1} = P^{-1}P = I$, and (18) and (19) imply that

$$M N^T = I - RS \qquad \text{and} \qquad P \Pi_1 = \Pi_2. \tag{20}$$

The new block diagonal matrix $\mathcal{T} = \text{diag}(\Pi_1, I, I)$ is defined and the transform of (16) using Π_1 and $\mathcal{T}$ implies

$$\Pi_1^T P \Pi_1 > 0, \qquad \mathcal{T}^T \Theta_j \mathcal{T} < 0, \qquad \text{for } j \in \{1, 2\}. \tag{21}$$

If we define the new variables as

$$\begin{aligned} \hat{A} &= N A_k M^T + N B_k C R + S B C_k M^T \\ &\quad + S(A + B D_k C) R \\ \hat{B} &= N B_k + S B D_k \\ \hat{C} &= C_k M^T + D_k C R \\ \hat{D} &= D_k \end{aligned} \tag{22}$$

then $F_j(X) = \mathcal{T}^T \Theta_j \mathcal{T}$ and $F_3(X) = \Pi_1^T P \Pi_1$ become LMIs as (4). ∎

Proof for (b): For full-order control design, M and N are invertible and (A_k, B_k, C_k, D_k) can be uniquely determined from ($\hat{A}, \hat{B}, \hat{C}, \hat{D}, R, S$) as

$$\begin{aligned} A_k &= N^{-1}(\hat{A} - (\hat{B} - S B \hat{D}) C R \\ &\quad - S B (\hat{C} - \hat{D} C R) - S(A + B \hat{D} C) R) M^{-T} \\ B_k &= N^{-1}(\hat{B} - S B \hat{D}) \\ C_k &= (\hat{C} - \hat{D} C R) M^{-T} \\ D_k &= \hat{D}. \end{aligned} \tag{23}$$

∎

REFERENCES

[1] H. Fillmore, I. Chasiotis, S. Cho, and G. Gillies, "Atomic force microscopy observations of tumour cell invadopodia: Novel cellular nanomorphologies on collagen substrates," *Nanotechnology*, vol. 14, pp. 73–76, 2003.

[2] A. Pantazi, A. Sebastian *et al.*, "Probe-based ultrahigh-density storage technology," *IBM J. Res. Develop.*, vol. 52, no. 4/5, pp. 493–511, Jul./Sep. 2008.

[3] D. Croft, G. Shedd, and S. Devasia, "Creep, hysteresis and vibration compensation for piezoactuators: Atomic force microscopy application," *Trans. ASME, J. Dyn. Syst., Meas. Control*, vol. 123, pp. 35–43, 2001.

[4] G. Schitter, K. Astrom, B. DeMartini, P. Thurner, K. Turner, and P. Hansma, "Design and modeling of a high-speed AFM-scanner," *IEEE Trans. Control Syst. Technol.*, vol. 15, no. 5, pp. 906–915, Sep. 2007.

[5] K. Leang and A. Fleming, "High-speed serial-kinematic AFM scanner: Design and drive considerations,," in *Proc. Amer. Control Conf.*, Jun. 2008, pp. 3188–3193.

[6] K. Leang and S. Devasia, "Hysteresis, creep and vibration compensation for piezoactuators: Feedback and feedforward control," in *Proc. 2nd IFAC Conf. Mechatronic Syst.*, 2002, pp. 283–289.

[7] Y. Wu and Q. Zou, "Iterative control approach to compensate for both the hysteresis and the dynamics effects of piezo actuators," *IEEE Trans. Control Syst. Technol.*, vol. 15, no. 5, pp. 936–944, Sep. 2007.

[8] A. J. Fleming and S. O. R. Moheimani, "A grounded-load charge amplifier for reducing hysteresis in piezoelectric tube scanners," *Rev. Sci. Instrum.*, vol. 76, no. 7, pp. 073707-1–073707-5, 2005.

[9] G. Schitter, P. Menold, H. F. Knapp, F. Allgower, and A. Stemmer, "High performance feedback for fast scanning atomic force microscopes," *Rev. Sci. Instrum.*, vol. 72, no. 8, pp. 3320–3327, Aug. 2001.

[10] S. Salapaka, A. Sebastian, J. P. Cleveland, and M. V. Salapaka, "High bandwidth nano-positioner: A robust control approach," *Rev. Sci. Instrum.*, vol. 73, no. 9, pp. 3232–3241, 2002.

[11] A. Sebastian and S. Salapaka, "Design methodologies for robust nano-positioning," *IEEE Trans. Control Syst. Technol.*, vol. 13, no. 6, pp. 868–876, Nov. 2005.

[12] S. Bashash and N. Jalili, "Robust adaptive control of coupled parallel piezo-flexural nanopositioning stages," *IEEE/ASME Trans. Mechatron.*, vol. 14, no. 1, pp. 11–20, Feb. 2009.

[13] J. Dong, S. Salapaka, and P. Ferreira, "Robust control of a parallel-kinematic nanopositioner," *Trans, ASME, J. Dyn. Syst., Meas. Control*, vol. 130, pp. 041007-1–041007-15, Jul. 2008.

[14] Q. Zou and S. Devasia, "Preview-based optimal inversion for output tracking: Application to scanning tunneling microscopy," *IEEE Trans. Control Syst. Technol.*, vol. 12, no. 3, pp. 375–386, May 2004.

[15] S. S. Aphale, S. Devasia, and S. O. R. Moheimani, "High-bandwidth control of a piezoelectric nanopositioning stage in the presence of plant uncertainties," *Nanotechnology*, vol. 19, no. 12, pp. 125503-1–125503-9, Mar. 2008.
[16] C. Lee and S. Salapaka, "Optimal model matching design for high bandwidth, high resolution positioning in AFM," in *Proc. 17th IFAC Conf.*, Seoul, Korea, Jul. 2008, pp. 9230–9235.
[17] C. Lee and S. M. Salapaka, "Robust broadband nanopositioning: Fundamental trade offs, analysis, and design in a two-degree-of-freedom control framework," *Nanotechnology*, vol. 20, no. 3, pp. 035501–035516, 2009.
[18] B. E. Helfrich, C. Lee, D. A. Bristow, X. Xiaohui, J. Dong, A. G. Alleyne, and S. Salapaka, "Combined $\mathcal{H}_\infty$-feedback and iterative learning control design with application to nanopositioning systems," in *Proc. Amer. Control Conf.*, Seattle, WA, Jun. 2008, pp. 3893–3900.
[19] K. Zhou, J. Doyle, and K. Glover, *Robust and Optimal Control*. Upper Saddle River, NJ: Prentice-Hall, 1996.
[20] M. Vidyasagar, *Control System Synthesis: A Factorization Approach*. Cambridge, MA: MIT Press, 1985.
[21] D. J. Hoyle, R. A. Hyde, and D. J. N. Limebeer, "An $\mathcal{H}_\infty$ approach to two degree of freedom design," in *Proc. IEEE Conf. Decision Control*, Dec. 1991, pp. 1581–1585.
[22] D. Shim, H. S. Lee, and L. Guo, "Mixed-objective optimization of a track-following controller using linear matrix inequalities," *IEEE/ASME Trans. Mechatron.*, vol. 9, no. 4, pp. 636–643, Dec. 2004.
[23] C. Du, L. Xie, J. N. Teoh, and G. Guo, "An improved mixed $\mathcal{H}_\infty/\mathcal{H}_2$ control design for hard disk drives," *IEEE Trans. Control Syst. Technol.*, vol. 13, no. 5, pp. 832–839, Sep. 2005.
[24] C. Scherer, P. Gahinet, and M. Chilali, "Multiobjective output-feedback control via LMI optimization," *IEEE Trans. Autom. Control*, vol. 42, no. 7, pp. 896–911, Jul. 1997.
[25] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control, Analysis and Design*. 2nd ed. New York: Wiley, 2005.
[26] G. E. Dullerud and F. Paganini, *A Course in Robust Control Theory : A Convex Approach* (Texts in Applied Mathematics 36). Berlin, Germany/New York: Springer-Verlag, 2000.
[27] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory* (Studies in Applied Mathematics 15). Philadelphia, PA: SIAM, Jun. 1994.
[28] M. de Oliveira, J. Geromel, and J. Bernussou, "An lmi optimization approach to multiobjective controller design for discrete-time systems," in *Proc. 38th IEEE Conf. Decision Control*, 1999, vol. 4, pp. 3611–3616.
[29] C. Schere, "Lower bounds in multi-objective $\mathcal{H}_\infty/\mathcal{H}_2$ problems," in *Proc. 38th IEEE Conf. Decision Control*, 1999, vol. 4, pp. 3605–3610.

Chibum Lee (S'08) was born in Seoul, Korea, in 1975. He received the B.S. and M.S. degrees in mechanical design and production engineering from Seoul National University, Seoul, Korea, in 1998 and 2000, respectively. He is currently working toward the Ph.D. degree in mechanical science and engineering at the University of Illinois at Urbana-Champaign, Urbana.

From 2000 to 2005, he was a Dynamics and Control Engineer with Hyundai Mobis, Seoul, Korea. His current research interests include application of control theory to nanoscale systems. He was engaged in the development of control design for nanopositioning systems and the new mode for atomic force microscopes.

Mr. Lee is a member of the American Society of Mechanical Engineers.

Srinivasa M. Salapaka (S'96–M'02) was born in Andhra Pradesh, India, in 1973. He received the B.Tech. degree in mechanical engineering from the Indian Institute of Technology Madras, Chennai, India, in 1995, and the M.S. and Ph.D. degrees in mechanical engineering from the University of California, Santa Barbara, in 1997 and 2002, respectively.

From 2002 to 2004, he was a Postdoctoral Associate in the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge. Since 2004, he has been with the Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana. His current research interests include controls for nanotechnology, combinatorial resource allocation, and numerical analysis of integral equations.

Dr. Salapaka received a National Science Foundation CAREER Award in 2005. He is a member of the American Society of Mechanical Engineers.