A robust loop-shaping approach to fast and accurate nanopositioning

Sumeet S. Aphale*, Antoine Ferreira and S. O. Reza Moheimani

* (Corresponding Author) Dr. Sumeet S. Aphale, School of Engineering, Kings College, University of Aberdeen, Aberdeen, UK. [s.aphale@abdn.ac.uk]
Prof. Antoine Ferreira, PRISME’ Institute, ENSI Bourges, France. [antoine.ferreira@ensi-bourges.fr]
Prof. S. O. Reza Moheimani, School of Electrical Engineering and Computer Science, University of Newcastle, Australia. [reza.moheimani@newcastle.edu.au]

Abstract

Inherent nonlinearities as well as their sharp lightly-damped resonant mode severely limit the positioning bandwidth of commercially available piezoelectric stack actuated nanopositioners. In this paper, we propose a hybrid control scheme that incorporates a $H_{\infty}$-based loop-shaping control scheme for resonance damping and an integral control for tracking. Experimental raster scans recorded in both open-loop and closed-loop operation are presented to show the substantial increase in positioning performance achieved by the implemented control scheme. It is also shown that the control scheme is robust under relatively large changes in platform resonance, an effect of platform loading.

Key words: Nanopositioning, $H_{\infty}$ loop-shaping, resonance damping, tracking, rastering

1. Introduction

The piezoelectric-stack actuated nanopositioning platform has emerged as a popular, commercially-available nanopositioning device due to its simple design, robust mechanical construction, ease of sensor integration, large motion ranges, almost flat motion profile and inherently low axial cross-coupling, [1, 2, 3, 4]. Most of these commercially available platforms are essentially parallel-kinematic manipulators with flexure guidance systems.
Figure 1 shows one such nanopositioner with integrated parallel-plate capacitive sensors for measuring axial displacement. These nanopositioners have a number of inherent performance-limiting factors such as lightly damped resonant modes, hysteresis and creep; that their accompanying rudimentary controllers, mostly PI or PID, cannot alleviate without drastically reducing their positioning bandwidth, [5, 6]. To push the boundaries of their positioning performance in terms of speed, linearity and accuracy, a number of evolved control algorithms have been proposed.

Inversion-based feedforward techniques for continuously tracking a desired trajectory have been investigated, [7, 8]. Inversion-based feedforward technique has also been applied to nanopositioning systems [9]. It was shown that by using this approach, it is possible to compensate for nonlinear effects such as creep and hysteresis [10, 11]. High sensitivity to modeling errors and plant parameter uncertainties are the main drawbacks of the feedforward control strategy. Feedback techniques that suppress the effects of uncertainties have been shown to improve the bandwidth of feedforward strategies [12]. Force-feedback for nano-manipulation was reported in [13]. In [14], robust tracking was reported. $H_\infty$ control algorithms for nano-scale tracking have also been investigated [15]. An in-depth overview of this field is given in
Variation in resonance frequency is one of the commonly occurring system uncertainties in a nanopositioning stage. These variations can occur due to a number of reasons such as a change in ambient temperature, humidity, atmospheric pressure and mechanical loading. The most effective method to make the system less sensitive to variations in resonance frequencies is to impart substantial damping to the dominant resonant mode of the nanopositioner using a feedback controller and ensure that the controller is sufficiently robust to changes in resonance frequency. Passive damping techniques such as shunts have been effective but need constant tuning and show a drastic performance degradation under variations in resonance frequencies. Adaptive shunts are needed to handle system uncertainties and have been reported in [18, 19]. Feedback controllers that impart substantial damping to the system have been documented over the years [20, 21, 22, 23, 24]. PPF control [25] and Resonant control [26] have shown substantial damping capability. Polynomial-based controller is another straightforward controller, [27], which is easy to construct for second-order plants, imparts substantial damping to the system, and is robust under resonance frequency variations [28, 29]. In recent years, Integral Resonant Control has also been applied successfully to damp the dominant resonant mode of custom-built and commercially available nanopositioners, [30, 31]. Combining such damping controllers that are insensitive to variations in resonance frequencies with a PI controller has been reported earlier [32], where it was noted that the high gain needed by the integrator to track a given trajectory limits the achievable scan speed using this technique to about 4 Hz. The need for high-speed high-resolution scanning at the nanometer scale has led to several control approaches being investigated, [33, 34]. Using the inversion-based feedforward technique in conjunction with the feedback, it was shown that the scan speed of the same commercially-available nanopositioner could be increased, [35].

In this paper, a $H_\infty$-based loop-shaping controller that robustly damps the platform resonance is combined with inversion-based feedforward and integral tracking control to ultimately result in fast and accurate nanopositioning. To further reduce the effects of hysteresis and creep, the fast axis of the nanopositioner is driven by charge rather than voltage using a custom-built charge amplifier, [36].
2. Experimental Setup

Figure 2 shows the experimental setup used for this work. The commercially available piezoelectric stack actuated nanopositioner used in this work is the PI-734 XY piezo-scanner that has a $100\mu m \times 100\mu m$ motion range (scan area). The E-509 sensor module was employed to procure displacement data from the in-built capacitive sensors of the nanopositioner. This module generates voltages proportional to the respective axis displacements. The control algorithm was implemented using the dSPACE DS1103 rapid prototyping system consisting of 16-bit A/D input channels and 16-bit D/A output channels operating in parallel at a sampling rate of 20 kHz. To reduce the effect of environmental exogenous vibrations and disturbances, the nanopositioner was mounted on a pneumatic vibration isolation table. All the transfer functions were recorded using a HP35670A Dynamic Signal Analyzer.

2.1. System Identification

The nanopositioning platform is treated as a two input - two output system, the inputs being the command signals for $x$ and $y$ axes and the outputs being the resulting displacements in each of the axes. The nanopositioner is driven by an amplifier using a swept-sine wave input generated by the HP35670A Dynamic Signal Analyzer. The motion of the nanopositioner in each axis is sensed by the two-plate capacitive sensor integrated into each
Figure 3: The frequency response of the two-input two-output nanopositioning platform measured from inputs (amplifier input voltage proportional to desired displacement) X and Y to sensor outputs (capacitive displacement sensor voltages proportional to resulting displacement). Note that the \( G_{xx}(s) \) has a resonant peak at 426 Hz, \( G_{yy}(s) \) has a resonant peak at 395 Hz and there is substantial decoupling (a minimum of \( \approx -30 \) dB between the two axes).
axis and this signal is then fed to the E-509 sensor module. Finally, the sensor module produces a voltage proportional to the platform displacement.

Figure 3 plots the four magnitude responses (from input X to output X \( \rightarrow G_{xx}(s) \), from input X to output Y \( \rightarrow G_{yx}(s) \), from input Y to output X \( \rightarrow G_{xy}(s) \) and from input Y to output Y \( \rightarrow G_{yy}(s) \)) for the nanopositioner. The first and dominant resonant peaks for the \( x \) and the \( y \) axes occur at 426 Hz and 395 Hz respectively. Throughout this work, we will address the axis with its first resonant peak at 426 Hz to be the fast axis and the one with its resonance at 395 Hz as the slow axis. As the most popular scan trajectory is a raster pattern employing a triangle wave (with high frequency components) on one axis and a slow ramp on the other, this is the scan trajectory used in our experimental analysis to test the positioning performance of the nanopositioner operated in closed-loop with the proposed control strategy. Consequently, the fast axis will be driven with a triangle wave input while the slow axis will move along a slow ramp / staircase trajectory. The fast axis of the nanopositioner was driven by a charge amplifier while the slow axis was driven by a voltage amplifier. Also note that though the cross-coupling is low throughout the frequency range, it is particularly high near the platform resonance. Damping the resonance modes of both the axes will also result in reduced cross-coupling, an extremely favorable property for accurate nanopositioning systems.

As operating point (D.C. Bias) as well as external loading tend to change the frequency response of the system, frequency responses were measured to ascertain the behavior of the nanopositioner to these operating condition variations. Figure 4(a) shows the frequency magnitude response of the nanopositioner for different operating points. As can be seen, variation in operating point leads to minor changes in the overall system gain that manifest clearly at lower frequencies. Change in loading conditions have a more pronounced impact of the frequency response of the nanopositioner. As shown in Figure 4(b), as the load increases, the resonance frequency as well as the magnitude of the resonance peak reduce.

3. Control Design

There are two main problems that have been targeted by the proposed control design viz: (i) the lightly-damped resonant mode of the nanopositioner and (ii) nonlinear effects such as hysteresis and creep. Additionally, as these nanopositioners are employed in applications that require the nanopo-
Figure 4: (a) Frequency responses of one axis of the nanopositioning platform at various operating points (resulting from changes in the DC bias applied to the piezoelectric actuator of the respective axis). (b) Frequency responses of one axis of the nanopositioning platform under different loading conditions.

3.1. $H_{\infty}$ Loop-shaping for robust resonance damping

We define the nominal model of the system to be controlled starting from the coprime factors on the left: $G(s) = \tilde{M}(s)^{-1}\tilde{N}(s)$. Then a perturbed model is written (see Fig. 5).

$$\tilde{G} = (\tilde{M} + \Delta_m)^{-1}(\tilde{N} + \Delta_n)$$

(1)

where $\tilde{G}$ is a left coprime factorization (LCF) of $G$, and $\Delta_m, \Delta_n$ are unknown and stable transfer functions representing the uncertainty. We can then define a family of models with the following expression:

$$\xi_\epsilon = \{ \tilde{G} = (\tilde{M} + \Delta_m)^{-1}(\tilde{N} + \Delta_n) : \| (\Delta_m \Delta_n) \|_\infty < \epsilon \}$$

(2)
where \( \epsilon_{\text{max}} \) represent the margin of maximum stability. The robust problem of stability is thus to find the greatest value of \( \epsilon = \epsilon_{\text{max}} \), such as all the models belonging to \( \xi_\epsilon \) can be stabilized by the same corrector \( K \). The problem of robust stability \( H_\infty \) amounts finding \( \gamma_{\text{min}} \) and \( K(s) \) stabilizing \( G(s) \) such as:

\[
\left\| \begin{pmatrix} I & K \end{pmatrix} (I - G K)^{-1} \begin{pmatrix} I & W_2 G W_1 \end{pmatrix} \right\|_\infty = \gamma_{\text{min}}^{-1} = \epsilon_{\text{max}}^{-1} = \epsilon_{\text{max}}^{-1} (3)
\]

However, [38] showed that the minimal value \( \gamma \) of is given by:

\[
\gamma_{\text{min}} = \epsilon_{\text{max}}^{-1} = \sqrt{1 + \lambda_{\text{sup}}(XY)}
\]

(4)

where \( \lambda_{\text{sup}} \) indicates the greatest eigenvalue of \( XY \). Moreover, for any value \( \epsilon < \epsilon_{\text{max}} \) a corrector stabilizing all the models belonging to \( \xi_\epsilon \) is given by:

\[
K(s) = B^T X (sI - A + BB^T X - \gamma^2 Z Y C^T C)^{-1} \gamma^2 Z Y C^T
\]

\[
Z = (I + Y X - \gamma^2 I)^{-1}
\]

\[
\gamma = \epsilon^{-1}
\]

(5)

where \( A, B \) and \( C \) are state matrices of the system defined by the function \( G \) and \( X, Y \) are the positive definite matrices and solution of the Ricatti equation.
3.1.1. The Loop Shaping Design Procedure

Contrary to the approach of Glover-Doyle, no weight function can be introduced into the problem. The adjustment of the performances is obtained by affecting an open modeling (Loop Shaping) process before calculating the compensator. The design procedure is as follows:

- We add to the matrix $G(s)$ of the system to be controlled, a pre-compensator $W_1$ and/or a post-compensator $W_2$. The nominal plant $G(s)$ and shaping functions $W_1$ and $W_2$ are combined in order to improve the performances of the system such as $G_a(s) = W_2(s)G(s)W_1(s)$ (see Fig. 6(a)).

- From coprime factorizations of $G_a(s)$, we apply the previous results to calculate $\epsilon_{\text{max}}$, and then synthesize a stabilizing controller $K$ ensuring a value of $\epsilon$ slightly lower than $\epsilon_{\text{max}}$.

- The final feedback controller is obtained by combining the $H_\infty$ controller $K(s)$ with the shaping functions $W_1$ and $W_2$ such that $G_a(s) = W_2(s)G(s)W_1(s)$ (See Figure 6(b)).

3.1.2. Robust Controller Synthesis

$H_\infty$ based loop-shaping control technique basically involves choosing suitable pre- and post-compensator weighting functions that shape the plant dy-
dynamics as desired. Once these weighting functions are chosen, an $H_\infty$-based controller $K(s)$ that robustly stabilizes the shaped plant is synthesized. For nanopositioning systems, it is apt to choose a pre-compensator to be a scalar gain that pushes the gain of the plant at low-frequencies to unity. Additionally, the post-compensator is chosen such that the higher-frequency dynamics (seen in the frequency response at frequencies greater than the first dominant resonant mode) roll-off at a faster rate. Consequently, for the nanopositioner used in this work, the pre-compensator weight $W_1$ is chosen to be 0.5 and the post-compensator $W_2$ is chosen to be

$$W_2 = \frac{5.053 \times 10^7}{s^2 + 4021s + 2.527 \times 10^7}.$$  \hspace{1cm} (6)$$

The $H_\infty$-based controller $K(s)$ that robustly stabilizes the shaped plant $G_a(s)$ is found to be:

$$K(s) = \frac{-2.358s^3 - 5735s^2 - 4.307 \times 10^7s + 1.021 \times 10^{11}}{s^3 + 1.243 \times 10^4s^2 + 9.15 \times 10^7s + 4.13 \times 10^{11}}$$ \hspace{1cm} (7)$$

This controller results in good gain and phase margins (G.M. = 10.3 dB at 872.17 Hz and P.M. = 68.8 deg at 552.26 Hz). Using a similar approach, the loop-shaping pre- and post-compensator weights as well as a $H_\infty$-based controller was designed for the slow axis of the nanopositioning platform. The details are given below.

$$W'_1 = 0.45$$

$$W'_2 = \frac{5.053 \times 10^7}{s^2 + 6032s + 2.527 \times 10^7}$$

$$K'(s) = \frac{-2.417s^3 - 2400s^2 - 1.46 \times 10^7s + 6.96 \times 10^{10}}{s^3 + 1.091 \times 10^4s^2 + 6.442 \times 10^7s + 2.385 \times 10^{11}}$$

Simulations predict almost a 22 dB damping augmentation at resonance for both the $x$ and $y$ axes of the nanopositioner.

3.2. Tracking control

To increase the accuracy of the nanopositioner as well as the positioning bandwidth, an integral tracking controller combined with the inversion-based feed-forward technique was implemented. The details of these controllers are given below.
3.2.1. Integral tracking

Due to the substantial damping of the resonant mode achieved by the $H_\infty$ loop-shaping controller discussed earlier, there is sufficient gain-margin in the system to implement a relatively high-gain integral tracking controller. Also, as the slow axis is given a ramp input, the integral tracking controller is ideally suited. Thus, a gain of 600 was selected for the implemented integrator. Figure 7 shows the overall control loop implemented for the slow axis.

3.2.2. Inversion-based Feed-forward control

Input-shaping using plant inversion is a well-known and popular technique for achieving accurate scan performance using nanopositioning devices [9, 35]. The inversion-based feedforward controller finds the input ‘$u$’ that, when fed to a linear system with known dynamics ‘$G_{xx}$’, produces the desired output ‘$y_d$’. In mathematical terms, the task is to find a $u$ such that

$$Y_d(i\omega) = G_{xx}(i\omega)U(i\omega)$$

where $U(i\omega)$ and $Y_d(i\omega)$ are Fourier transforms of $u$ and the desired trajectory $y_d$ respectively. In most scanning applications, the desired trajectory is a periodic triangle waveform and can be expressed as a Fourier series viz:

$$y_d(t) = \sum_{k=1}^{\infty} A_k \sin(\omega_k t)$$

where $A_k = \frac{8}{\pi^2 k^2} \sin\left(\frac{\pi k}{2}\right)$ and $\omega_k = 2\pi k f$ with $f$ being the fundamental frequency of the triangular waveform. Therefore, the Fourier transform $Y_d(i\omega)$
is a sum of impulses with magnitude $A_k$. Hence, an input $u(t)$ based on the inverted system dynamics $\frac{1}{G_{xx}(i\omega_k)}$, given by

$$u(t) = \sum_{k=1}^{\infty} \frac{A_k}{|G_{xx}(i\omega_k)|} \sin(\omega t - \phi_k)$$

where

$$\phi_k = \angle G_{xx}(i\omega_k), \quad k = 1, 2, \ldots, \quad (11)$$

should result in $y_d(t)$ being the output, for more details refer to [39]. However, this technique suffers from two poignant drawbacks viz:

1. The desired trajectory $y_d(t)$, and hence $u(t)$, have an infinite number of harmonics and are consequently not bandlimited.

2. The construction of $u(t)$ depends on the accuracy of the model $G_{xx}(i\omega)$

Thus, an inversion-based feedforward controller cannot achieve satisfactory tracking performance in the presence of plant uncertainties. Variations in resonance frequency and damping are two important uncertainties that can occur in nanopositioning platforms. In the presence of such uncertainties, the achievable positioning bandwidth of the platform is substantially degraded. In the proposed input-shaping method using plant inversion, the closed-loop plant, $G_{cl}$, is identified and inverted. The desired input, $u_{ref}$, is fed to this inverted plant model, $\tilde{G}_{cl}^{-1}$, and this output, $u_{inv}$, is fed to the actual closed-loop plant, $G_{cl}$, to obtain the desired output $y_{des}$. Note that the closed-loop plant has been identified for a finite number of frequencies and thus errors due to higher unmodeled input harmonics are present but are very small. A pictorial representation of this scheme is given in Figure 8(a). Figure 8(b) shows an example of the generated inversion-based input. Note that the high-frequency components of the triangle wave (close to the tip) have been amplified accordingly to compensate for the closed-loop system gain at those particular frequencies.

4. Experimental Results and Analysis

Open- and closed-loop frequency responses of the system are plotted in Figure 9. It is clear that the cross-coupling is also reduced in closed-loop. It
Figure 8: (a) Inversion-based feed-forward strategy implemented on the fast axis of the nanopositioner for high accuracy and bandwidth. (b) Inversion-based input ($u_{inv}$) for a desired output ($y_{des}$) of a 20 Hz triangle trajectory.

is also clear that tracking bandwidth of the slow axis is limited to 60 Hz due to the presence of the integral loop action. To check for robustness under variation in resonance frequency, the platform was loaded with a mass of 50 gm. This shifts the resonance by a substantial amount, from 426 Hz to 350 Hz. The implemented control strategy is capable of damping the resonant mode by more than 20 dB, even under the presence of resonance frequency uncertainties, see Figure 10.

Open and closed loop plots for 10 Hz, 20 Hz and 40 Hz triangular trajectories are recorded. These are plotted in Figure 11. The open- and closed-loop RMS errors are tabulated in Table 1. The scan range is limited to 16 $\mu$m due to the current limit of the charge drive. As can be seen from Table 1, the RMS error is maximum at 40 Hz and is equal to 0.24 % of the total scan range (in closed-loop).

<table>
<thead>
<tr>
<th>Fundamental frequency for the triangular waveform (Hz)</th>
<th>Open-loop RMS error (nanometers)</th>
<th>Closed-loop RMS error (nanometers)</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>81.8</td>
<td>21.4</td>
</tr>
<tr>
<td>20</td>
<td>224.4</td>
<td>32.9</td>
</tr>
<tr>
<td>40</td>
<td>653.7</td>
<td>38.8</td>
</tr>
</tbody>
</table>

Table 1: Table showing the RMS value of error between the desired and measured scans at three fundamental frequencies.

Raster scans were obtained by exciting the fast axis with a periodic triangle wave and the slow axis with a synchronized pseudo-ramp input signal.
Figure 9: The fast axis (X) has loop-shaping control. The slow axis (Y) has loop-shaping control and integral tracking. The fast axis is charge driven and thus the nonlinear effects of hysteresis are minimal. The slow axis is driven by voltage and the integral control is implemented to enforce tracking and reduce nonlinear effects. The closed-loop 3-dB bandwidth of the slow axis is 60 Hz.

Raster scans were recorded at 10 Hz, 20 Hz and 40 Hz as shown in Figure 12. 15 µm wide scans were obtained with a line spacing of 72 nm. At 10 Hz, the scans have a full 15 µm linear range. At 20 Hz, the linear range drops to 80 % of the full range and at 40 Hz, the linear range is greatly reduced to less than 55 %. This is due to the tracking bandwidth limitation of the slow axis (3 dB at 60 Hz). The noise is predominantly quantization noise coupled with noise from the capacitive sensor and has an RMS value of about 13 nm.
Figure 10: Open- and closed-loop frequency responses of the nominal and the system loaded with a mass of 50 gm. As the loading increases, the bandwidth of the closed-loop system reduces drastically.

5. Conclusion

An integrated control strategy for commercially available piezoelectric nanopositioners that incorporates both damping control and trajectory tracking was proposed. The damping control was based on the $H_\infty$ loop-shaping algorithm and was shown to be robust under significant variations in resonance frequency. The damped resonant mode facilitates the implementation of a high-gain integral controller for trajectory tracking. This proposed strategy was tested on the PI-734 piezo-scanner and the performance as well as the robustness of the proposed control scheme was validated. Raster scans were obtained at frequencies as high as 40 Hz. It was seen that the performance degrades gracefully even with wide variations in resonance frequency.
Figure 11: (a), (c) and (e) show the open-loop (undamped) plots for 10 Hz, 20 Hz and 40 Hz triangular scan outputs respectively. (b), (d) and (f) show the closed-loop (feedback and inversion-based input shaping combined) tracking results for the same. Note that in closed-loop, the tracking is visibly superior. As the frequency content in the scan signal increases, the scan performance of the closed-loop worsens due to loss of bandwidth.
Figure 12: Raster scans obtained at 10 Hz, 20 Hz and 40 Hz. The solid green line denotes the measured data while the dotted blue line is the reference. Note that as the scanning frequency increases, the effective usable scan area decreases. This is because the sharp direction shifts that occur during a raster scan cannot be accurately followed by the band-limited tracking loop.

References


