Control Methodologies for Precision Positioning Systems

Xu Chen and Masayoshi Tomizuka

Abstract— **We have many servo systems that require nano/micro level positioning accuracy. This requirement sets a number of interesting challenges from the viewpoint of sensing, actuation, and control algorithms. This paper considers the control aspect for precision positioning. In motion control systems at nano/micro levels, it has been noted that the control algorithm should be customized as much as possible to the spectrum of disturbances. We will examine how prior knowledge about the disturbance spectrum should be utilized in the design of control algorithms and what are advantages of such prior knowledge. The algorithms will be evaluated on a simulated hard disk drive (HDD) benchmark problem and a laboratory setup for a wafer scanner that is equipped with a laser interferometer for position measurement.**

Index Terms— **precision control, feedback design, Youla parameterization, disturbance rejection, loop shaping, digital control**

I. Introduction

A precision servo system aims at accurately positioning the controlled object(s) to follow the desired trajectories. The controlled object here, for instance, can be a read/write head for accessing data in a commercial hard disk drive (HDD), or a stage carrying wafers (called wafer stage) in a wafer scanner in the semiconductor industry. In both examples, the required position accuracies are quite high. In the HDD example, the scenario can be mimicked by imagining an airplane flying at 5,000,000 mph above a 100,000-lane highway, to follow the center of a lane whose width is only fraction of an inch [1]! The error upper bound for the wafer-stage case is even smaller, nowadays about one magnitude lower than that of the HDDs. Such ultra-high precision is achieved by careful consideration of various disciplines in mechanical engineering. From the viewpoint of system integration, we can classify the design elements to the following four categories:

- *hardware and sensing components*: we may want to choose fluid or air bearings for reduced friction, laser interferometers or high-precision encoders for accurate measurements, and/or piezo-electronic/MEMS actuators for fine positioning;
- *operation environment*: such as friction-isolation tables, clean room, and thermostatic chambers;
- *task plans and arrangements*: such as well designed trajectories, and the arrangement of task repetitions in a manufacturing process;
- *servo control algorithms*–the software development to operate the actuators and correct the accumulated errors in the previous three design processes.

A well-designed precision system reflects optimized considerations in all the above categories. Servo control, as the finalizing step, is responsible to synthesize and compensate as much as possible the imperfections from previous design elements. Such imperfections are often inevitable. To list a few, we have:

- *hardware imperfection* which comes from system resonances, imperfections in bearings and gears, torque ripples in motors, periodic disturbance from cooling fans, delays in motor drivers and signal acquisition, etc;
- *environmental disturbance* such as turbulent airflow in small-scale fast-speed servo, structural vibrations that are extremely difficult to mechanically model (e.g., vibrations induced from high-power audio speakers in HDDs), and resonances of the operation table;
- *special errors due to the task nature*, such as repeated trajectories in various industrial manufacturing processes.

Errors induced from the above sources place great challenge in reaching position accuracies at the micro/nano scale. Fortunately, part of them–such as imperfect motor rotation, repeated trajectories, and fan noises–are repeatable once the hardware and the trajectory are fixed. Other errors–such as those caused by (environmental) vibrations–although may vary case by case, are at least structured, and can be compensated by carefully designed servo controllers.

In this tutorial paper we consider the control aspects of precision positioning systems. Overall, feedforward control has been generally accepted to be necessary for accurate trajectory tracking. Further, if a process is repeatedly performed for multiple iterations, learning control such as repetitive control and iterative learning control can be quite effective in reducing the errors in the future iterations/repetitions. These methods can be understood as to perform error compensation in the iteration domain.[1] Customized feedback control on the other hand can reduce the errors within each iteration, or the full time scale when the process is not repeatable. In the presence of various aforementioned error sources, the feedback loop in a precision position-

The authors are with Department of Mechanical Engineering, University of California, Berkeley, CA, 94720, USA (email: maxchen@me.berkeley.edu; tomizuka@me.berkeley.edu)

[1]Although working in the iteration domain, repetitive control is a feedback algorithm.

ing system needs to have the flexibility of providing different closed-loop features for error reduction, and the reliability of strong stability under different loop modifications.

In the remainder of the paper we will unfold these basic design concepts and discuss detailed algorithms to realize them. In Section II we study the identification of the problem from a spectral perspective. We focus in Section III on feedback algorithms including learning control concepts for precision servo, and provide several comments on feedforward designs in Section IV. Section V concludes the paper.

Throughout the paper we will use results from several simulated benchmark problems and laboratory hardware setups.

II. Identification of the (structural) problem

A primal assumption we made is that the position error has structured characteristics. Otherwise, if errors are pure white noises, little can be done for servo improvement. Structured errors here can be classified from the time- or frequency-domain perspective. As an example, we discuss next the wafer-scanning process in lithography. This is one key step for circuit fabrication in the semiconductor industry. To print the circuit, the wafer is exposed to patterned ultraviolet lights which come through a mask carried by a reticle stage. The wafer stage and the reticle stage move the wafer and the mask in a synchronized manner. Due to limited size of the lens, only a small part of the wafer is exposed at each scan, and the wafer is moved from one field to another in between the scans. The scanning process is repeated until all required areas on the wafer have been exposed under the light.

Fig. 1 illustrates an experimental setup of the key wafer-scanner components at the Mechanical Systems Control Laboratory, UC Berkeley. It uses linear permanent magnet motors as actuators and laser interferometers to measure the position of the two stages. Fig. 2 shows the movement of the reticle stage during a scaled scanning process. The measurement here is obtained at a sampling time of $T_s = 0.0004$ sec. This is the baseline result with an PID controller in the closed loop. No feedforward algorithms have been applied yet. From the repetitive nature of the process, if we append all errors at continuous iterations, the position error in Fig. 2 will be replicated to yield a time sequence in Fig. 3. The Fast Fourier Transform (FFT) of this error vector is shown in Fig. 4, where we observe strong energy components at the multiples of a fundamental frequency. Figs. 2-4 analyze just one simple example encountered in applications. However, regardless of the shape of the trajectory, the periodic error structure would not change. The fundamental frequency can be analytically obtained by $f_o = 1/T_{\text{trajectory}}$, where $T_{\text{trajectory}}$ is the period of one iteration. This is a direct result of the Fourier-series theory. In the example in Figs. 2-4, $f_o = 1/(300 \times 0.0004) \approx 8.33$ Hz, which can be seen to match the FFT result in Fig. 4.

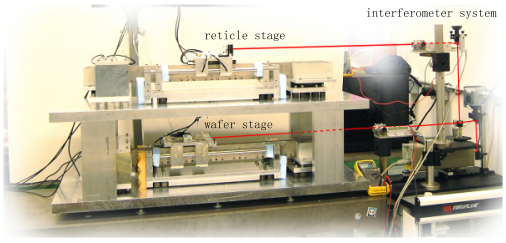

Fig. 1. An experimental testbed of wafer scanners

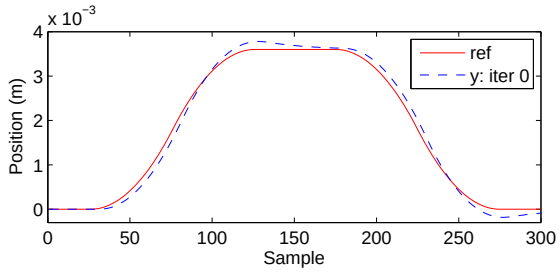

Fig. 2. An example scanning trajectory

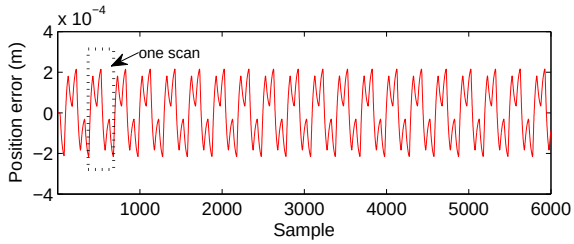

Fig. 3. Errors in a repeated scanning process

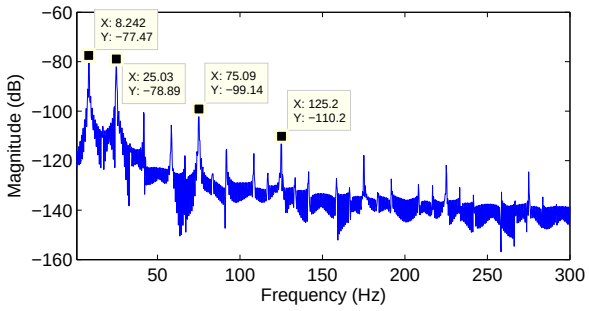

Fig. 4. FFT of the error in Fig. 3

Fig. 5 shows a section of the experimental results when we applied a longer reference trajectory. The control of the position errors is particularly important when the stage is moving at a constant speed. Actual wafer scanning occurs here. A zoomed-in view at the constant-speed region indicates that small ripples exist

in the position error. Such errors are caused by imperfect interactions between magnetic fields and conductors in the linear motor. The ripple structure is more clearly explained in Fig. 6, where we have plotted the spectrum of the position error collected at the constant-speed regions. The peak at around 18 Hz is the main contributor to the ripples in Fig. 5 (we can compute the period of the ripple in the time domain and confirm the frequency).

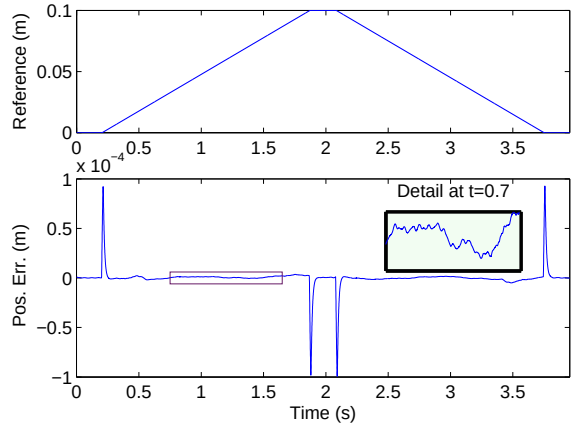

Fig. 5. Time-domain tracking result for a longer scanning process

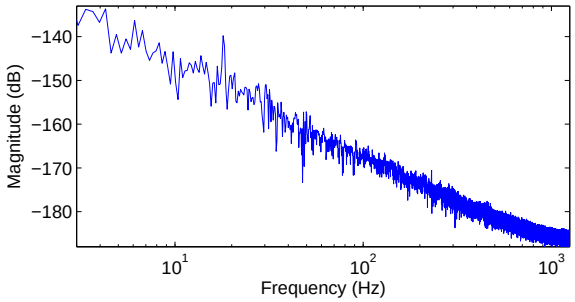

Fig. 6. FFT of the errors under the constant-speed scanning process

The above is a spectral perspective of a reference-tracking example. Similar observation can be made for regulation control. Essentially, at high-precision scales, the structure of small errors can usually be better revealed in the frequency domain. The situation can be much more complicated than the one shown in Fig. 6 (see e.g. Fig. 12). Yet the idea of reducing the peaks in the error spectra is common in precision servo.

III. Feedback controller parametrization

A precision positioning system is designed to have an accurate plant dynamics P, which is usually linear and time-invariant. If we focus on using linear controllers (at least at the steady state), then position servo design is essentially about loop shaping. In this section, we first briefly review several fundamental concepts, then provide algorithms for flexible designs to address the structured errors discussed in the previous section.

Consider the block diagram of a standard feedback design in Fig. 7. The closed-loop transfer function from the disturbance d_o to the plant output y, is defined as the sensitivity function $S \triangleq 1/(1+PC)$. The complementary sensitivity function is given by $T \triangleq 1-S$, which equals $PC/(1+PC)$, i.e., T is the transfer function from the reference y_d to y. The sensitivity function measures the closed-loop disturbance-attenuation property, while the complementary sensitivity function defines how the system responds to the reference input as well as the sensor noise. Consider a typical magnitude response of S in Fig. 8. Below the bandwidth ω_c, the magnitude of $S(j\omega)$ is less than 1, hence the attenuation of d_o in Fig. 7. Wherever the magnitude of $S(j\omega)$ is small, $T(j\omega)$ will be close to unity due to the fundamental relationship $S+T=1$. This means that at frequencies where we achieve good disturbance rejection, we also obtain improved reference tracking.

The bandwidth in Fig. 8 can not be pushed to be arbitrarily large. From the practical perspective, any mechanical system would not be able to respond to arbitrarily fast control inputs, due to hardware (such as motors, gears, etc) limitations. It is common practice to keep the gain of the controller small at high frequencies. From the theoretical perspective, under mild conditions,[2] it is inevitable to have $|S(j\omega')|>1$ at certain frequencies if $|S(j\omega)|<1$ holds over some frequency interval, namely, when certain disturbance components are attenuated, some other disturbance components will be amplified. This is the "waterbed" effect which comes from Bode's Integral Theorem, a fundamental result of linear control design.

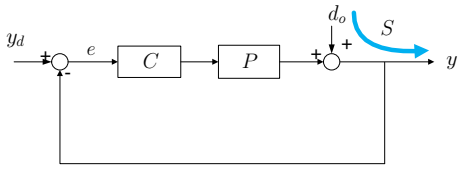

Fig. 7. A standard closed-loop system under feedback control

The preceding discussion serves as a baseline for precision positioning control. Well formulated tools such as PID, lead-lag, and H_∞ controllers are available to achieve a loop shape that is similar to the one in Fig. 8. We focus next on how we can perform safe and flexible add-on modifications to a standard loop shape.

Based on Fig. 7, consider a new controller

$$\tilde{C} = \frac{C+Q}{1-PQ} \tag{1}$$

[2]For continuous-time systems, the waterbed effect holds if the relative degree of the loop transfer function $L=PC$ is no less than 2. In the discrete-time case, waterbed always holds.

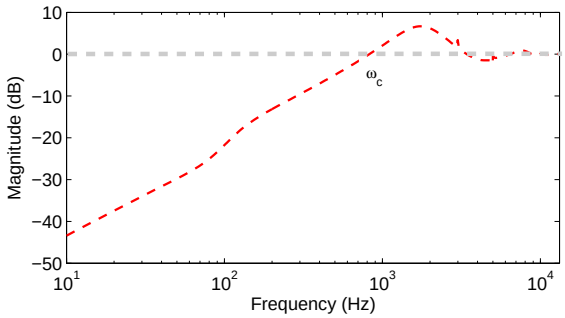

Fig. 8. A typical magnitude response of the sensitivity function

where Q is a stable proper rational transfer function to be discussed shortly. The new sensitivity and complementary sensitivity functions are then given by

$$\tilde{S} = \frac{1}{1+P\tilde{C}} = \frac{1-PQ}{1+PC} \tag{2}$$

$$\tilde{T} = 1-\tilde{S} = \frac{PC-PQ}{1+PC}. \tag{3}$$

$\tilde{S}$ and $\tilde{T}$ are affine functions of Q. When Q and P are both stable, $\tilde{S}$ and $\tilde{T}$ are also stable, as $1/(1+PC)$ and $PC/(1+PC)$ are the original sensitivity and complementary sensitivity functions. It is not difficult to further verify that the new closed-loop system is stable as well. Hence $\tilde{C}$ is a stabilizing controller for the plant P. Actually, when P and C are stable, *any* feedback controller that stabilizes P can be expressed in the form of (1), by choosing a proper Q that is stable and rational. This is the result of the simplest form of Youla parameterization [2]. Besides the appealing feature of being able to express any stabilizing controllers, Youla parameterization has another useful property: $\tilde{S}$ in (2) is decomposed to be the product of the original sensitivity function $1/(1+PC)$ and the add-on element $1-PQ$. Since stability has already been guaranteed, the affine form of Q in $\tilde{S}$ makes it quite easy for designers to directly impact the sensitivity function.

The general Youla parameterization for single-input-single-output systems is as follows: define the set $\mathcal{S}$ = {stable, proper, rational transfer functions}. If a plant $P = N/D$ can be stabilized by a negative-feedback controller $C = X/Y$, with (N, D) and (X, Y) being coprime pairs over $\mathcal{S}$, then any stabilizing feedback controller can be parameterized as

$$C_{all} = \frac{X+DQ}{Y-NQ}: \ Q \in \mathcal{S}, \ Y(\infty) - N(\infty)Q(\infty) \neq 0. \tag{4}$$

Here a pair (N,D) is called coprime over $\mathcal{S}$ if $U \in \mathcal{S}$, $V \in \mathcal{S}$ and there exists $U \in \mathcal{S}$, $V \in \mathcal{S}$ such that $UN+VD=1$.[3]

[3]The inequality $Y(\infty) - N(\infty)Q(\infty) \neq 0$ makes the closed loop to be well-posed. This condition is usually easy to satisfy for practical problems. Consider, for example, the case where $P = z^{-1}$ and $C = 0.8$. We have $N = z^{-1}$, $D = 1$, $X = 0.8$, $Y = 1$, and $Y(z=\infty) - N(z=\infty)Q(z=\infty) = 1$.

After simplification, the new sensitivity function with controller (4) is

$$\tilde{S} = \frac{1}{1+PC_{all}} = \frac{1}{1+PC}\left[1 - \frac{N}{Y}Q\right] \tag{5}$$

which is again, affine in Q.

It can now be seen that, in the choice of (1), the plant and the baseline controller are parameterized by $P = P/1$ and $C = C/1$, namely, $N = P$, $D = 1$, $X = C$, and $Y = 1$ in (4). These are valid coprime factorizations when P and C are stable, since we can choose, e.g., $U = C/(1+PC) \in \mathcal{S}$ and $V = 1/(1+PC) \in \mathcal{S}$ to make $UN+VD = 1$. Such choices of N, D, X, and Y make (2) a simplified (and easier-to-use) version of (5).

The remaining design task about choosing Q in (2) depends on the desired closed-loop response and the transfer function P. Candidate Q-filter designs can be found, for example, where a linear combination of some basis transfer functions (e.g., $\sum_{i=0}^{k_Q} \theta_i z^{-i}$ in discrete-time schemes) is used to form Q, and adaptive/H_∞ control is applied to find the scaling coefficients for the combination. Regardless of the tools that are used, to make $\tilde{S}$ in (2) small at certain frequencies, we require PQ to be close to 1 in the same region. Consider a block-diagram realization in Fig. 9. In the inner loop marked by $\star$, Q should then approximate P^{-1} at the desired frequencies. As $1-PQ$ becomes small, the overall controller (1) will then have high gain (increased control efforts) in the interested frequency regions.

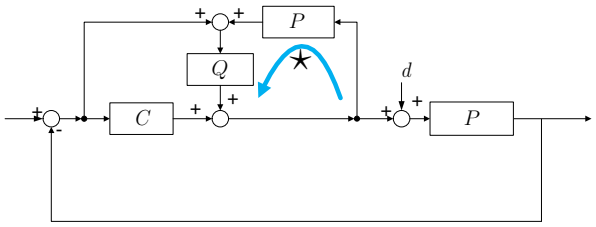

Fig. 9. A forward-model Youla parameterization

Instead of a direct inversion of P by Q, we discuss below an alternative scheme that will later bring additional simplifications to (2). Although a general analysis can be made, we focus on precision servo systems and assume that the baseline controller C is stable.

Consider designing Youla parameterization in two steps: first to perform an explicit stable model inversion P^{-1}, and then to use Q for controlling the amount of loop shaping and disturbance rejection. Of course, a perfect inversion of the plant dynamics is practically not possible, and we have to adopt a nominal version $\hat{P}^{-1}$. In addition, during discrete-time implementation, the sampled-data transfer function $\hat{P}^{-1}(z^{-1})$ is possibly not causal/realizable, since a practical plant P usually has delays (suppose there are m steps of them) that become z^m when inverted. For realizability, we consider

using instead $z^{-m}\hat{P}^{-1}(z^{-1})$. For the moment we assume $\hat{P}^{-1}(z^{-1})$ is stable. This is not difficult to achieve for precision positioning systems. We will come back to this point later in the article. A block diagram of the aforementioned idea is shown in Fig. 10. Compared to Fig. 9, this is a discrete-time scheme that is closer to actual implementation, and modifications have been made to reduce the dependence of the plant information in the ($\star$) loop in Fig. 9.

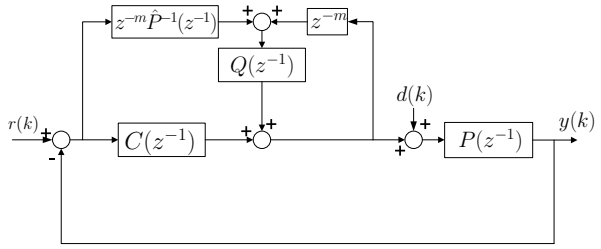

Fig. 10. A discrete-time inverse-based Youla parameterization

The transfer function of the overall feedback controller in Fig. 10 is

$$C_{all}(z^{-1}) = \frac{C(z^{-1}) + z^{-m}\hat{P}^{-1}(z^{-1})Q(z^{-1})}{1 - z^{-m}Q(z^{-1})} \tag{6}$$

In (6), high-gain control is directly provided by $1/(1 - z^{-m}Q(z^{-1}))$, instead of by $1/(1-PQ)$ in (1). This reduced dependence on P is also reflected in the frequency sensitivity function $\tilde{S}(e^{-j\omega}) = \tilde{S}(z^{-1})\big|_{z=e^{j\omega}}$. If $\hat{P}(e^{-j\omega}) = P(e^{-j\omega})$, (6) gives

$$\tilde{S}(e^{-j\omega}) = \frac{1}{1 + P(e^{-j\omega})C_{all}(e^{-j\omega})} = \frac{1 - e^{-mj\omega}Q(e^{-j\omega})}{1 + P(e^{-j\omega})C(e^{-j\omega})} \tag{7}$$

$$= S_0(z^{-1})(1 - z^{-m}Q(z^{-1}))\big|_{z=e^{j\omega}} \tag{8}$$

From (8), the add-on design narrows down to designing the term $1 - z^{-m}Q(z^{-1})$, which is simpler than $1 - PQ$ in (2) and $1 - NQ/Y$ in (5). Independent from the plant dynamics, several general design guidances for $Q(z^{-1})$ can now be made.

A. Low-frequency servo enhancement

An intuitive choice for $Q(z^{-1})$ is a low-pass filter. In this case, if the delay of the plant is not very large, then the frequency transfer function $1 - e^{-j\omega m}Q(e^{-j\omega})$ would be small when $Q(e^{-j\omega})$ is approximately one at low frequencies, and close to unity when $|Q(e^{-j\omega})| \ll 1$ at high frequencies. A direct result is that any bias disturbance will be rejected. This is because under the low-pass assumption we have $1 - e^{-j\omega\times 0}Q(e^{-j\times 0}) = 1 - Q(e^{-j\times 0}) = 0$ at $\omega = 0$. Therefore an integral action is built into the closed-loop controller.

B. Narrow-band disturbance rejection

Vibrations are frequency-dependent signals by nature. Since the closed-loop bandwidth can not be arbitrarily increased, vibrations at frequencies above the servo bandwidth are fundamentally more difficult to handle. Actually, such band-limited disturbances, if strong enough, will significantly limit the servo performance even when their frequencies are below the bandwidth. To compensate such vibrations, we would need some special feedback adjustment. In the meantime, we commonly want to keep as much as possible the original loop shape, which has been achieved via careful baseline design. A candidate add-on design is to use the proposed Youla parameterization and design a notch shape for $1 - z^{-m}Q(z^{-1})$ such as the one shown in the bottom plot of Fig. 11. To demonstrate the flexibility, we introduce five notches in the magnitude of $1 - z^{-m}Q(z^{-1})$. The first three notches are very close to each other, the other two are separated at higher frequencies. When $1 - e^{-mj\omega}Q(e^{-j\omega})$ is close to unity, $\tilde{S}(e^{-j\omega}) = S_0(z^{-1})(1 - z^{-m}Q(z^{-1}))\big|_{z=e^{j\omega}}$ will be close to $S_0(e^{-j\omega})$, hence the maintenance of the original loop shape. At frequencies where $1 - z^{-m}Q(z^{-1})$ has very small gains in Fig. 11, $\tilde{S}(e^{-j\omega})$ will also be very small, hence the attenuation of vibrations. Such loop modifications will preserve the original servo bandwidth and provide additional disturbance-attenuation capacity at the center frequencies of $1 - z^{-m}Q(z^{-1})$.

An example about disturbance attenuation in a HDD benchmark problem [3] is presented in Fig. 12. This benchmark has been used in a number of publications on information storage systems. We implemented the compensation scheme for rejecting two strong vibrations at around 1100 Hz and 1500 Hz. Both vibrations occur at frequencies above the baseline servo bandwidth (1060 Hz), and not attenuated in the standard feedback setting. The depth of attenuation is quite strong. The two originally sharp spectral peaks are actually visually not detectable after compensation. This is due to the deep notch shape of $1 - z^{-m}Q(z^{-1})$ in Fig. 11.

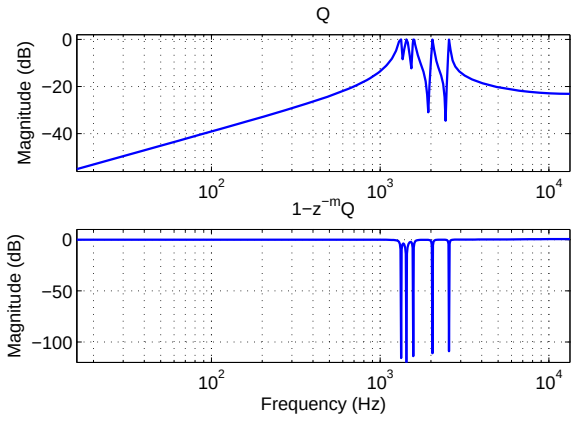

Fig. 11. A Q-design example for narrow-band disturbance rejection

Take the example of $m = 1$. The key concept of Q design is as follows: it is desired to have a notch

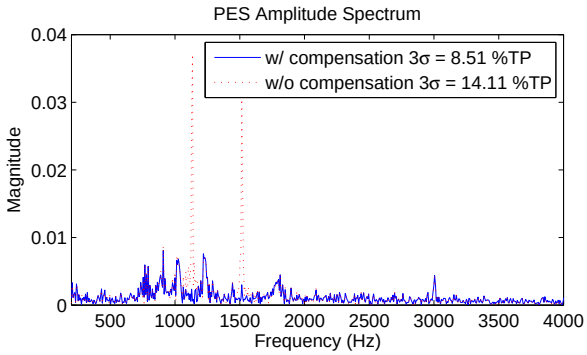

Fig. 12. Frequency spectra of the position error signals in a simulated HDD benchmark: 1 TP (Track Pitch) = 254nm in this example.

shape for $1-z^{-1}Q(z^{-1})$. Designing $1-z^{-1}Q(z^{-1}) = \frac{1-2\cos\omega_0 z^{-1}+z^{-2}}{1-2\alpha\cos\omega_0 z^{-1}+\alpha^2 z^{-2}}$ and solving for $Q(z^{-1})$ gives

$$Q(z^{-1}) = \frac{(\alpha^2-1-a^2(\alpha-1))-(\alpha-1)az^{-1}}{1+a\alpha z^{-1}+\alpha^2 z^{-2}},\ a=-2\cos(\omega_0) \tag{9}$$

Extending (9), if we want multiple notches, letting $Q(z^{-1}) = B_Q(z^{-1})/A_Q(z^{-1})$ with $1-z^{-1}Q(z^{-1}) = A(z^{-1})/A(\alpha z^{-1})$, $A(\alpha z^{-1}) = \prod_{i=1}^{n}(1-2\cos(2\pi\Omega_i T_s)\alpha z^{-1}+\alpha^2 z^{-2})$, and $A(z^{-1}) = A(\alpha z^{-1})\big|_{\alpha=1}$, we can obtain the general solution:

$$A_Q(z^{-1}) = 1+\sum_{i=1}^{n-1} a_i(\alpha^i q^{-i}+\alpha^{2n-i}q^{-2n+i})+a_n\alpha^n q^{-n}+\alpha^{2n}q^{-2n}$$

$$B_Q(q^{-1}) = \sum_{i=1}^{2n}(\alpha^i-1)a_i q^{-i+1},\ a_i = a_{2n-i} \tag{10}$$

Here $\alpha(<1)$ controls the width of the attenuation regions, and can be taken to be very close to 1 for narrow-band disturbance rejection. The coefficients a_i's in (9) and (10) determine the center frequencies of $Q(z^{-1})$. In (9), $a=-2\cos(2\pi\Omega_{Hz}T_s)$ where Ω_{Hz} is the desired notch frequency. In (10), a_i and Ω_i are connected by the mapping $A_Q(z^{-1}) = A(\alpha z^{-1})$.

When $m>1$, several additional steps are needed since $1-z^{-m}Q(z^{-1}) = A(z^{-1})/A(\alpha z^{-1})$ would not have a realizable solution. An approach that uses Diophantine identity to address this issue is provided in [4].

C. Repetitive control

This is for addressing the problem discussed in Figs. 2-4. An intuition about repetitive control is that, if the same type of disturbance occurs after a fixed period of time, i.e., $(1-z^{-N})d(k)=0$ where N is the period of the disturbance, then at the next occurrence of $d(k)$ we can learn and reduce the resulting error, no matter how it behaves within one period of time. From the spectral perspective, the magnitude response of $1-z^{-m}Q(z^{-1})$ in Fig. 13 provides an example loop shape for the sensitivity function: at multiples of the fundamental frequency, $|1-e^{-jm\omega}Q(e^{-j\omega})|$ gives low gains to $\tilde{S}(e^{-j\omega})$. Meanwhile at other frequencies $|1-e^{-jm\omega}Q(e^{-j\omega})|$ is approximately unity, yielding no change to $|\tilde{S}(e^{-j\omega})|$. The Q filter to achieve Fig. 13 is

$$Q(z^{-1}) = \frac{(1-\alpha^N)z^{-(N-m-n_q)}}{1-\alpha^N z^{-N}} z^{-n_q} q(z,z^{-1}) \tag{11}$$

where N and m are as defined previously; $q(z,z^{-1})$ is a zero-phase low-pass filter with order n_q; and α controls the "waterbed" effect by the following theorem:

Theorem 1: [5], [6] When $P(e^{-j\omega}) = \hat{P}(e^{-j\omega})$, the worst-case amplification of the non-repetitive disturbance is $(2/(1+\alpha^N)-1)\times 100\%$ for repetitive control (RC) using (11). In conventional RC, the amplification is at most 100%, which corresponds to the case with $\alpha=0$ in (11).

An α closer to 1 gives reduced amplification of the non-repetitive disturbance. A trade off in this case is that the algorithm requires more accurate knowledge about the period of the disturbance/trajectory.

Fig. 14 shows the experimental result of the algorithm to reduce the errors mentioned in Fig. 3. With the repetitive learning control, the errors are observed to have reduced to two-magnitude lower than the case without compensation.

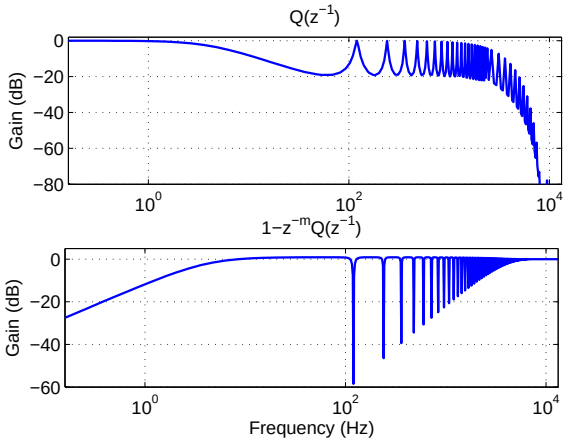

Fig. 13. A Q-design example for repetitive control

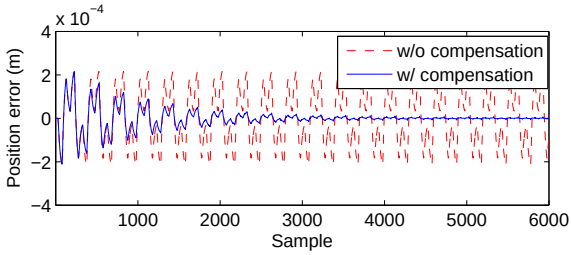

Fig. 14. Tracking errors with the Q-filter configuration in (11).

D. General band-limited vibration compensation

The loop shaping in Section III-B was for narrow-band disturbance rejection. When excitation sources are rich in frequency, a wider attenuation bandwidth

is needed in $1-z^{-m}Q(z^{-1})$. This can be achieved by changing the design parameter α in (9)-(10), which is related to the width of the -3dB passband for $1-z^{-m}Q(z^{-1})$ by $BW \approx (1-\alpha^2)/[(\alpha^2+1)(\pi T_s)]$ (Hz) [7]. As we push further the disturbance attenuation, the trade-off amplification due to the waterbed effect will also increase. Additional design considerations are therefore necessary for safe implementation. Fig. 15 shows the loop-shaping result on a HDD benchmark example. No large sensitivity amplifications are observed due to: (i), a gain scheduling in the Q filter with $Q_{\text{implement}} = k_Q Q(z^{-1})$, $k_Q \le 1$ and (ii), the use of a damped notch filter $F_{nf}(z^{-1}) = A(\beta z^{-1})/A(\alpha z^{-1})$ ($\alpha < \beta < 1$) in the design of $Q(z^{-1})$. Auxiliary zeros and convex optimization can be applied for specific magnitude constraints on $Q(z^{-1})$ [8].

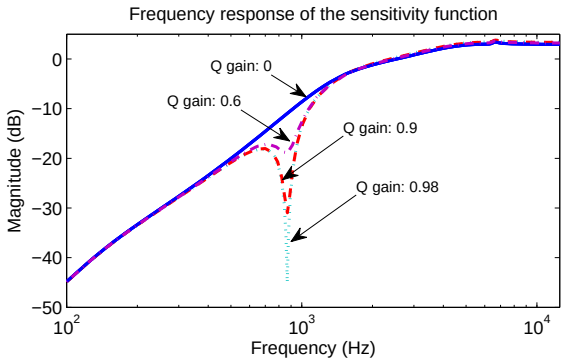

Fig. 15. The control of attenuation efforts via gain scheduling on Q

IV. Remarks

We have discussed the general design concepts about loop shaping and add-on Youla parameterization. In this section, we provide several notes that are relevant to practical implementation.

The plant model and its inverse: the use of the (inverse) model information has been essential in the discussed control schemes. Such a model is usually readily available in industries, as it not only provides convenience for feedback design, but also is beneficial for tasks such as fault detection and disturbance isolation. By the physical construction, the input to actuators in a precision positioning system is usually a voltage/current signal that is approximately linear w.r.t. motor torque, and the measurement is commonly an angular or linear position. The inverse system dynamics thus usually has a double-differentiator type of frequency response due to Newton's law. It is thus not difficult to obtain a stable nominal $\hat{P}^{-1}(z^{-1})$ that accurately presents the low-frequency dynamics.[4] Direct differentiation at high frequencies is not practical and there are high-frequency resonances. High-frequency unstable zeros, if any, can be addressed by e.g., shifting them to be inside the unit circle when forming of $\hat{P}^{-1}(z^{-1})$.

Robust stability: strictly speaking, after we put $\hat{P}$ in the controller design, we are implementing a robust version of Youla parameterization. Notice that in the examples in Figs. 11 and 13, the Q filters are designed such that their magnitudes are small at high frequencies, namely, we are not changing the loop behavior at ultra-high frequencies. This is important for implementation, as it is practically not possible to have an exact mathematical model for mechanical systems. A robust stability analysis can give us upper bounds of the plant uncertainties for the closed-loop system to maintain stable [5]. Intuitively, flexible loop shaping can be readily achieved at frequency regions where the model $\hat{P}^{-1}(z^{-1})$ accurately reflects the actual system dynamics; at high frequencies where the plant behavior itself can not be well predicted, the magnitude of $Q(z^{-1})$ (and hence the add-on control efforts) is recommended to be kept small for robust stability.

Feedforward control: we have discussed the feedback perspective for precision servo. Due to space limit, detailed feedforward (FF) discussions are not included in this tutorial paper. Necessity of FF control comes from the simple fact that feedback designs have bandwidth limitations. In the FF class of control algorithms, model-based design is also of essential importance. Inverse complementary sensitivity and inverse plant dynamics are two common approaches for FF design. When the process is repetitive, iterative learning control (ILC) is another powerful tool for error correction in the iteration domain. An introduction and some case studies are provided during the actual conference presentation.

Block-diagram intuition: for regulation control, the output of $Q(z^{-1})$ in Fig. 10 actually approximates $-d(k)$, i.e., the negative of the input disturbance. Hence the structure has the property of an input-disturbance observer. Discussions on this point are provided in [5].

Adaptive configuration: when the disturbance spectra is not known (but with known structure), adaptive control can be applied to update the parameters of $Q(z^{-1})$ online. Investigation for the case of narrow-band disturbance rejection is provided in [4], [9].

Reference of Youla parameterization: for additional materials about Youla parameterization, readers can refer to, for example, [10], [11]. In the discrete-time case, the simplest example to design Q is to use $Q(z^{-1}) = \theta_0 + \theta_1 z^{-1} + \theta_2 z^{-2} + ... + \theta_{k_Q} z^{-k_Q}$, namely, a finite impulse response (FIR) filter, and form parameter adaptation algorithms to find a set of θ_i to minimize the feedback error. This idea is also used in adaptive inverse control. A rich reference in this area is [12].

V. Conclusion

In this paper, we have studied the control methodologies for precision positioning systems. From the view points of hardware, operation environment, and

[4]Friction can also be approximated by a second-order mode with damping and spring.

task arrangements, we discussed common issues that may arise to control design engineers, and analyzed the error structures from a spectral perspective. The central concept we try to deliver is that add-on loop shaping is a convenient tool for addressing various servo problems, and we used multiple simulation and experimental results on actual engineering problems to validate the presented designs. Besides the examples used in this article, the control structure has also been tested under other systems such as electrical power steering in automotive vehicles and active suspension systems. A special session about vibration rejection on the latter system is to be presented in the 2013 European Control Conference, where different control algorithms, including the one in this paper, are applied to the same system and compared in various aspects such as computation time, steady-state and transient performances.

VI. Acknowledgment

We gratefully acknowledge the hardware support from Nikon Research Corporation of America, National Instruments, and Agilent Technologies.

References

[1] A. Al Mamun, G. Guo, and C. Bi, *Hard disk drive: mechatronics and control*. CRC Press, 2007.

[2] D. Youla, H. Jabr, and J. Bongiorno Jr, "Modern Wiener-Hopf design of optimal controllers Part II: The multivariable case," *IEEE Trans. Autom. Control*, vol. 21, no. 3, pp. 319–338, 1976.

[3] IEEJ, Technical Commitee for Novel Nanoscale Servo Control, "NSS benchmark problem of hard disk drive systems," http://mizugaki.iis.u-tokyo.ac.jp/nss/, 2007.

[4] X. Chen and M. Tomizuka, "A minimum parameter adaptive approach for rejecting multiple narrow-band disturbances with application to hard disk drives," *IEEE Trans. Control Syst. Technol.*, vol. 20, no. 2, pp. 408 –415, march 2012.

[5] ——, "New repetitive control with improved steady-state performance and accelerated transient," *to appear in IEEE Trans. Control Syst. Technol.*, 2013.

[6] M. Tomizuka, T.-C. Tsao, and K.-K. Chew, "Analysis and synthesis of discrete-time repetitive controllers," *ASME Journal of Dynamic Systems, Measurement, and Control*, vol. 111, no. 3, pp. 353–358, 1989.

[7] X. Chen, A. Oshima, and M. Tomizuka, "Inverse based local loop shaping for vibration rejection in precision motion control," in *The 6th IFAC Symposium on Mechatronic Systems (to appear), Hangzhou, China, April 10-12*, 2013.

[8] X. Chen and M. Tomizuka, "Decoupled disturbance observers for dual-input-single-output systems with application to vibration rejection in dual-stage hard disk drives," in *Proc. 2012 ASME Dynamic Systems and Control Conf., and 2012 Motion & Vibration Conf.*, 2012, pp. 1544–1554.

[9] ——, "Unknown multiple narrow-band disturbance rejection in hard disk drives–an adaptive notch filter and perfect disturbance observer approach," in *Proc. 2010 ASME Dynamic Systems and Control Conf.*, 2010, pp. 936–970.

[10] J. C. Doyle, B. A. Francis, and A. Tannenbaum, *Feedback control theory*. Macmillan, 1992, vol. 134.

[11] B. D. Anderson, "From youla-kucera to identification, adaptive and nonlinear control," *Automatica*, vol. 34, no. 12, pp. 1485 – 1506, 1998.

[12] B. Widrow and E. Walach, *Adaptive Inverse Control, Reissue Edition: A Signal Processing Approach*. Wiley-IEEE Press, Nov. 2007.