DATA-DRIVEN CONTROL STRATEGY FOR A RETICLE STAGE IN A LITHOGRAPHIC TOOL

Yi Jiang¹, Kaiming Yang¹, Yu Zhu¹ and Chuxiong Hu¹
¹Department of Mechanical Engineering
Tsinghua University
Beijing, China

INTRODUCTION
Data-driven control is used to address the practical difficulty on obtaining accurate models. The rationale is that the controller is directly generated from the IO collection data, without the need of dynamics modeling. Most data-driven methods solve the design problem by iteratively minimizing an $H_2$ criterion, and the iterative law is [1]

$$p^{k+1} = p^k - \gamma^k R^{-1} \nabla J |_{p^k}$$

where $p$ is the controller parameter, $\nabla J$ is the gradient, $k$ is the iteration number, $R$ is a positive definite matrix and $y$ is the step size.

In this paper, a data-driven control strategy is proposed to synthetically build the control system of a 6-DOF ultra-precision dual-stroke stage. The stage, namely reticle stage which is a critical lithographic component, moves the pattern in synchronization with the wafer.

DESIGN AND CONTROL CONCEPT
The developed reticle stage consists of a short-stroke (SS) stage and a long-stroke (LS) stage. Figure 1 shows the structural configuration. SS stage is driven by voice coil motors (VCMs) to achieve contactless 6-DOF movement. Magnetic gravity compensators (MGCs) are mounted to counteract the gravity. The position is measured by laser interferometers with 0.6 nm resolution. Two LS stages are symmetrically distributed on the two sides of SS stage. LS stage is driven by an ironless linear motor and the measurement system includes eddy sensor and linear encoder.

FIGURE 1. (a) The schematic and; (b) The real assembly of the dual-stroke reticle stage.

Figure 2 shows the control concept. Under the diagonal dominance assumption that mostly can be achieved by static input dynamic decoupling $\Psi$, the system combines multi-SISO feedback control $C$ and decoupling feedforward control $F$.

FIGURE 2. Control scheme of the system
For simplicity, only SS stage is considered in this paper. As high-order and uncertainty characteristics, a data-driven strategy is proposed with an integrated design for $\Psi_s$, $C_s$ and $F_s$.

STATIC INPUT DYNAMIC DECOUPLING $\Psi_s$
The global force $F = [F_x,F_y,F_z,T_x,T_y,T_z]^T$, the local force $f = [f_x,f_y,f_z,f_{cy},f_{cz},f_{cx},f_{cy},f_{cz}]^T$ and the control force $F_c = [F_{cx},F_{cy},F_{cz},T_{cx},T_{cy},T_{cz}]^T$ satisfy

$$F = (\Psi'_s)^{-1} f = (\Psi'_s)^{-1} \Psi'_s F_c$$

where $\Psi'_s$ is the optimal value of $\Psi_s$. Due to the inevitable modeling error, the offset $p = [p_x, p_y, p_z]$ between the actual and theoretical center of mass (COM) will render coupling, as shown in figure 3.

FIGURE 3. Force acted on the SS stage
This paper shows a data-driven method to deal with the coupling caused by the unknown offset. As additional torque will be created, in a transitional movement, the optimization argument is

$$\arg\min_{b} \left\| A \begin{bmatrix} p_x & p_y & p_z \end{bmatrix} - b \right\|_k$$

$$A = \begin{bmatrix} F_{cy} - F_{cz} & F_{cz} - F_{cy} & F_{cy} - F_{cx} \end{bmatrix}$$

$$b = T_{cx} + T_{cy} + T_{cz}$$
Combining (1) and (3) gives the update law
\[ P^{k+1} = P^k - \gamma^k \left( A^T A \right)^{-1} A^T b \]  
(4)

**MULTI-SISO FEEDBACK CONTROLLER \( C_s \)**

With the diagonal dominance achieved by \( \Psi_s \), the multi-SISO feedback controller \( C_s \) is then to be designed. We define \( y_e \) as the desired response, and objective function is

\[ J(p) = \frac{1}{N} \sum_{n=0}^{N-1} \left[ \hat{y}(n) \right]^2 = \frac{1}{N} \sum_{n=0}^{N-1} \left[ (y - y_e) \right]^2 \]  
(5)

To obtain the ideal parameter \( p_0 \), a direct data-driven method is proposed, which highlights its convergence performance. Assuming that reference \( r \) is the step signal, \( J(p) \) can be transferred to the discrete-frequency domain as

\[ J(p) = \frac{1}{2N} \sum_{k=-N}^{N} \left| T(k,p) - T_d(k) \right|^2 P_r(k) \]  
(6)

where \( \hat{r} \) is \( r \) padded with zeros to length \( 2N-1 \) and \( P_r(k) \) denotes the power spectrum. The gradient of \( J(p) \) can be deduced as

\[ \nabla J = \frac{1}{N} \sum_{k=-N}^{N} P_r \left| S(p) \right|^2 \left\{ S^* S(p) \right\} \left( p - p_0 \right) \]  
(7)

\[ = M(p)(p - p_0) \]

With (7), a novel iterative update law is proposed as

\[ P^{k+1} = P^k - \gamma^k M^{-1} \nabla J_{|p^k} \]  
(8)

Furthermore, an unbiased and simplified approximation method of \( M \) and \( \nabla J \) is proposed, which will be discussed as follows.

**MIMO FEEDFORWARD CONTROLLER \( F_s \)**

With the feedback loop tuning completed, tracking performance is further improved by the feedforward control. Considering the servo error is the crucial performance, \( J(p) \) is chosen as

\[ J(p) = e(p)^T e(p) \]  
(9)

where \( e(p) = [e_x e_y e_z e_{x_0} e_{y_0} e_{z_0}] \) is the overall error. And Gauss-Newton algorithm is employed

\[ \nabla J_{|p^k} = 2 \nabla e(p)^T e(p) \]  
(10)

This paper develops an unbiased and simplified approximation of \( \nabla e \) based on Toeplitz matrix, which can also be used to estimate \( M \) and \( \nabla J \). For a discrete time-invariant system with zero initial state, the output \( y(n) \) can be defined as

\[
\begin{bmatrix}
  y(0) \\
  y(1) \\
  \vdots \\
  y(N-1)
\end{bmatrix} = \begin{bmatrix}
  D & 0 & \ldots & 0 \\
  CB & D & \ldots & 0 \\
  \vdots & \vdots & \ddots & \vdots \\
  CA^{N-2}B & \ldots & CB & D
\end{bmatrix} \begin{bmatrix}
  u(0) \\
  u(1) \\
  \vdots \\
  u(N-1)
\end{bmatrix}
\]

where \( \Omega \) is the Toeplitz matrix consisting of Markov parameters. \( \nabla e \) can be deduced as

\[ \nabla e(p) = \begin{bmatrix}
  T_{xx} & \ldots & T_{xz} \\
  \vdots & \ddots & \vdots \\
  T_{tx} & \ldots & T_{zz}
\end{bmatrix} \begin{bmatrix}
  \Pi & 0 \\
  \vdots & \Pi
\end{bmatrix} = T_{\text{step}} \gamma \]  
(12)

where \( \gamma \) depends on the reference \( r \) and the structure of \( F_s \), which are known for the designer. And \( T_{ij} \) is the map between \( j \)-direction process input \( u_i \) and \( i \)-direction servo error \( e_j \), which can be identified by the impulse response trial.

**EXPERIMENTS AND CONCLUSIONS**

The validity of the data-driven strategy has been verified by three experiments, as shown in Figure 4. The comparison between before and after optimization indicates that the performance is improved remarkably under the proposed tuning strategy. These comparative experiments definitely demonstrate the validity and effectiveness of the proposed data-driven tuning strategy.

![Figure 4. Related experiments: (a) optimize \( \Psi_s \) on a planar stage; (b) optimize \( C_s \) on a cable stage; (c) and (d) optimize \( F_s \) on a wafer stage.](image)

**ACKNOWLEDGEMENTS**

This work was supported in part by the National Nature Science Foundation of China under Grant 51105226.

**REFERENCES**