Predictive control of thermally induced wavefront aberrations

A. Haber,[1,*] A. Polo,[2] I. Maj,[1] S.F. Pereira[2], H.P. Urbach,[2] and M. Verhaegen[1]

[1]*Delft Center for Systems and Control, Delft University of Technology, Mekelweg 2, 2628 CD Delft, The Netherlands*
[2]*Optics Research Group, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands*

**a.haber@tudelft.nl*

Abstract: In this paper we experimentally demonstrate the proof of concept for predictive control of thermally induced wavefront aberrations in optical systems. On the basis of the model of thermally induced wavefront aberrations and using only past wavefront measurements, the proposed adaptive optics controller is able to predict and to compensate the future aberrations. Furthermore, the proposed controller is able to correct wavefront aberrations even when some parameters of the prediction model are unknown. The proposed control strategy can be used in high power optical systems, such as optical lithography machines, where the predictive correction of thermally induced wavefront aberrations is a crucial issue.

OCIS codes: (110.1080) Active or adaptive optics; (110.5220) Photolithography; (220.1000) Aberration compensation; (350.6830) Thermal lensing.

References and links

1. S. J. Sheldon, L. V. Knight, and J. M. Thorne, "Laser-induced thermal lens effect: a new theoretical model," Appl. Opt. **21**, 1663–1669 (1982).
2. P. Hello and J.-Y. Vinet, "Analytical models of transient thermoelastic deformations of mirrors heated by high power cw laser beams," J. Phys. France **51**, 2243–2261 (1990).
3. R. Lawrence, D. Ottaway, M. Zucker, and P. Fritschel, "Active correction of thermal lensing through external radiative thermal actuation," Opt. Lett. **29**, 2635–2637 (2004).
4. M. Kasprzack, B. Canuel, F. Cavalier, R. Day, E. Genin, J. Marque, D. Sentenac, and G. Vajente, "Performance of a thermally deformable mirror for correction of low-order aberrations in laser beams," Appl. Opt. **52**, 2909–2916 (2013).
5. B. Canuel, R. Day, E. Genin, P. La Penna, and J. Marque, "Wavefront aberration compensation with a thermally deformable mirror," Classical Quant. Grav. **29**, 085012 (2012).
6. S. Piehler, C. Thiel, A. Voss, M. Abdou Ahmed, and T. Graf, "Self-compensation of thermal lensing in optics for high-brightness solid-state lasers," Proc. SPIE **8239**, 82390Z–82390Z–9 (2012).
7. H. Haferkamp and D. Seebaum, "Beam delivery by adaptive optics for material processing applications using high-power co2 lasers," Proc. SPIE **2207**, 156–164 (1994).
8. J. Bekaert, L. Van Look, G. Vandenberghe, P. Van Adrichem, M. Maslow, J.-W. Gemmink, H. Cao, S. Hunsche, J. Neumann, and A. Wolf, "Characterization and control of dynamic lens heating effects under high volume manufacturing conditions," Proc. SPIE **7973**, Optical Microlithography XXIV, 79730V–79730V (2011).
9. D. H. Beak, J. P. Choi, T. Park, Y. S. Nam, Y. S. Kang, C.-H. Park, K.-Y. Park, C.-H. Ryu, W. Huang, and K.-H. Baik, "Lens heating impact analysis and controls for critical device layers by computational method," Proc. SPIE **8683**, Optical Microlithography XXVI, 86831Q–86831Q–8 (2013).
10. S. Halle, M. Crouse, A. Jiang, Y. van Dommelen, T. Brunner, B. Minghetti, M. Colburn, and Y. Zhang, "Lens heating challenges for negative tone develop layers with freeform illumination: a comparative study of experimental vs. simulated results," Proc. SPIE **8326**, 832607–832607–19 (2012).

#194428 - $15.00 USD Received 22 Jul 2013; accepted 19 Aug 2013; published 5 Sep 2013

11. C. Bikcora, M. Van Veelen, S. Weiland, and W. Coene, "Lens heating induced aberration prediction via nonlinear kalman filters," IEEE Trans. Semicond. Manuf. **25**, 384–393 (2012).
12. F. Staals, A. Andryzhyieuskaya, H. Bakker, M. Beems, J. Finders, T. Hollink, J. Mulkens, A. Nachtwein, R. Willekers, P. Engblom, T. Grunner, and Y. Zhang, "Advanced wavefront engineering for improved imaging and overlay applications on a 1.35 na immersion scanner," Proc. SPIE **7973**, 79731G-79731G-13 (2011).
13. K. Liu, Y. Li, F. Zhang, and M. Fan, "Transient thermal and structural deformation and its impact on optical performance of projection optics for extreme ultraviolet lithography," Jpn. J. Appl. Phys **46**, 6568–6572 (2007).
14. Y. Li, K. Ota, and K. Murakami, "Thermal and structural deformation and its impact on optical performance of projection optics for extreme ultraviolet lithography," J. Vac. Sci. Tech. B **21**, 127–129 (2003).
15. P. A. Spence, S. E. Gianoulakis, C. D. Moen, M. Kanouff, A. Fisher, and A. K. Ray-Chaudhuri, "System performance modeling of extreme ultraviolet lithographic thermal issues," J. Vac. Sci. Tech. B **17**, 3034–3038 (1999).
16. H. Lück, K.-O. Müller, P. Aufmuth, and K. Danzmann, "Correction of wavefront distortions by means of thermally adaptive optics," Opt. Commun. **175**, 275 – 287 (2000).
17. M. A. Arain, W. Z. Korth, L. F. Williams, R. M. Martin, G. Mueller, D. B. Tanner, and D. Reitze, "Adaptive control of modal properties of optical beams using photothermal effects," Opt. Express **18**, 2767–2781 (2010).
18. S. Ravensbergen, P. Rosielle, and M. Steinbuch, "Deformable mirrors with thermo-mechanical actuators for extreme ultraviolet lithography: Design, realization and validation," Precis. Eng. **37**, 353–363 (2013).
19. S. Ravensbergen, "Adaptive optics for extreme ultraviolet lithography : actuator design and validation for deformable mirror concepts," Ph.D. thesis, Technische Universiteit Eindhoven (2012).
20. F. Roddier, *Adaptive optics in astronomy* (Cambridge University Press, 1999).
21. R. Tyson, *Principles of Adaptive Optics, Third Edition*, Series in Optics and Optoelectronics Series (CRC Press-INC, 2010).
22. C. Mack, *Fundamental Principles of Optical Lithography: The Science of Microfabrication* (Wiley, 2008).
23. W. Nowacki, *Dynamic Problems of Thermoelasticity* (Springer, 1975).
24. M. Verhaegen and V. Verdult, *Filtering and System Identification: A Least Squares Approach* (Cambridge University Press, 2007).
25. Adaptica Srl, "Saturn user manual," http://www.adaptica.com/site/en/pages/saturn .
26. A. Haber, A. Polo, S. K. Ravensbergen, H. P. Urbach, and M. Verhaegen, "Identification of a dynamical model of a thermally actuated deformable mirror," Opt. Lett. **38**, 3061–3064 (2013).
27. D. Malacara, *Optical shop testing* (Wiley-Interscience, 2007).
28. E. Fernandez and P. Artal, "Membrane deformable mirror for adaptive optics: performance limits in visual optics." Opt. Express **11**, 1056–1069 (2003).
29. S. Bonora and L. Poletto, "Push-pull membrane mirrors for adaptive optics." Opt. Express **14**, 11935–11944 (2006).
30. A. Polo, A. Haber, S. F. Pereira, M. Verhaegen, and H. P. Urbach, "An innovative and efficient method to control the shape of push-pull membrane deformable mirror," Opt. Express **20**, 27922–27932 (2012).
31. A. Haber, A. Polo, C. S. Smith, S. F. Pereira, P. Urbach, and M. Verhaegen, "Iterative learning control of a membrane deformable mirror for optimal wavefront correction," Appl. Opt. **52**, 2363–2373 (2013).
32. M. Phan, R. Lim, and R. Longman, "Unifying Input-Output and State-Space Perspectives of Predictive Control," Princeton University, Department of Mechanical and Aerospace Engineering Technical Report No. 3044 (1998).
33. A. Haber and M. Verhaegen, "Moving horizon estimation for large-scale interconnected systems," in press, IEEE Trans. Autom. Control (2013).
34. G. Vdovin and M. Loktev, "Deformable mirror with thermal actuators," Opt. Lett. **27**, 677–679 (2002).

1. Introduction

In high power optical systems almost each element absorbs a portion of the beam's energy. The absorbed energy creates thermoelastic deformations and variation of refractive index of optical elements [1–3]. Consequently, it induces wavefront aberrations in the system. In this paper, the aberrations caused by heating of optical elements will be called *the Thermally Induced Wavefront Aberrations* (TIWA).

TIWA can limit the performance of a large variety of high power optical systems. For example, in gravitational wave interferometers high power lasers induce aberrations that can significantly decrease the sensitivity of the instruments [2, 4, 5]. TIWA can also degrade the beam quality of the lasers used in material processing [6, 7]. Furthermore, due to a constant demand for higher productivity and less production costs, the power transmitted through the projection optics of lithography machines constantly increases. Consequently, the energy absorbed by projection optics induces wavefront aberrations that can compromise the resolution of the

#194428 - $15.00 USD Received 22 Jul 2013; accepted 19 Aug 2013; published 5 Sep 2013

system [8–12]. In the next generation of optical lithography machines, that will use Extreme UltraViolet (EUV) sources, degradation of resolution due to the heating of optical elements will become even more severe [13–15].

Several types of active optical devices and adaptive optics (AO) concepts have been proposed for compensation of thermally induced aberrations in gravitational wave detectors [3–5, 16, 17]. In optical lithography machines, active optical elements have been introduced for correction of wavefront aberrations [8, 12]. Furthermore, in the next generation of lithographic machines (EUV lithographic - EUVL machines), deformable mirrors (DM)s will be used for wavefront correction [18, 19]. However, correction of TIWA in lithographic machines might not be possible using the classical AO techniques [20, 21]. This is because the classical AO techniques require that a wavefront sensor (WFS) and a DM are connected using a real-time feedback. This condition is not fulfilled in lithography machines because wavefront aberrations can be measured only before and after the exposure of certain number of wafers [8]. Furthermore, because the measurement time decreases the wafer throughput, the total measurement time should be as short as possible. Ideally, the wavefront should be measured only at the beginning of the exposure process and after exposure conditions change. Obviously, the classical feedback control algorithms, on which most of the AO techniques rely upon, cannot be applied in this scenario. Hence, novel AO control algorithms need to be developed. These new AO control algorithms should be able to predict the future behavior of the wavefront aberrations and to compensate them using only past measurement data.

In this paper we experimentally demonstrate the proof of concept for predictive control of thermally induced aberrations. On the basis of the model of TIWA and using only past wavefront measurements, the proposed controller is able to predict and to compensate the future wavefront aberrations. Furthermore, the proposed AO controller is able to correct wavefront aberrations even when some parameters of the prediction model are unknown. Beside optical lithography, the proposed predictive controller can be used in other high power optical systems where it is not possible to establish a real-time feedback between the controller and the wavefront sensor.

This paper is organized as follows. In Section 2, we present a problem formulation and we describe an AO experimental setup. In Section 3 and Section 4, we develop the predictive control algorithm and present the experimental results, respectively. In Section 5, the conclusions are drawn.

2. Problem description and experimental setup

Optical lithography is a technology that uses electromagnetic radiation to project mask patterns onto a photo-resist on a semiconductor wafer. The main components of a lithographic machine are: source, illumination optics, reticle stage (with a mask), projection optics and the wafer stage [22].

Because it is not possible to establish a real-time feedback between a wavefront sensor and a controller, the correction of TIWA in a lithography machine should be done in a predictive manner. Namely, the controller should be able to anticipate the future wavefront aberrations and to correct them. Obviously, the anticipation of the future wavefront aberrations must be done on the basis of the model of the TIWA. Furthermore, in order to do accurate prediction, the controller also needs to use past wavefront measurements (the main reason why the controller needs past wavefront measurements for prediction will be explained in Section 3).

The model of TIWA describes how exposure conditions influence the dynamical behavior of the wavefront aberrations. The exposure conditions that dominantly influence the wavefront aberrations are numerical aperture, source shape, reticle and mask pattern diffraction, exposure dose, throughput and resist stack [8]. In this paper these exposure conditions will be called

#194428 - $15.00 USD Received 22 Jul 2013; accepted 19 Aug 2013; published 5 Sep 2013

the inputs of the TIWA model. The TIWA model consists of the two main parts [8]. The first part relates the inputs with the distribution of exposure energy (intensity distribution) over the surfaces of the optical elements. This relation is established by first computing the full mask diffraction orders and then convolving them with the illumination source to obtain the diffraction pattern [8]. The computed diffraction pattern determines the intensity distribution on the optical elements. The second part consists of a thermoelastic system of Partial Differential Equations (PDEs), which relates the intensity distribution with thermoelastic deformations of the optical elements [2, 23].

The model of TIWA can be obtained using two approaches. The first approach relies on first principles modeling. For example, the TIWA model can be derived by finite element discretization of the thermoelastic system of PDEs. The second modeling approach is to identify the model directly from experimental data [24]. Identification can be performed on the basis of the wavefront measurements collected during the testing and calibration of a lithographic machine. However, during the exposure of the wafers, the exposure conditions usually differ from the exposure conditions used in machine calibration and testing. For example, different types of masks can be used during the exposure of wafers and their patterns can significantly differ from the patterns of the masks used in machine testing. In mathematical terms this means that during the wafer exposure, inputs of the TIWA model are usually unknown and they need to be determined. There are two ways to overcome this problem. If the new exposure conditions are precisely known then the inputs can be recalculated and the model can be updated. For example, if a new mask is projected, then it is necessary to calculate new intensity distributions on the optical elements. These calculations can be performed using the knowledge of the mask's geometry [8]. However, this is a computationally challenging problem that needs to be solved in real-time [22]. Another approach, that we propose in this paper, is to estimate unknown inputs or to estimate intensity distribution on optical elements using measured wavefront aberrations.

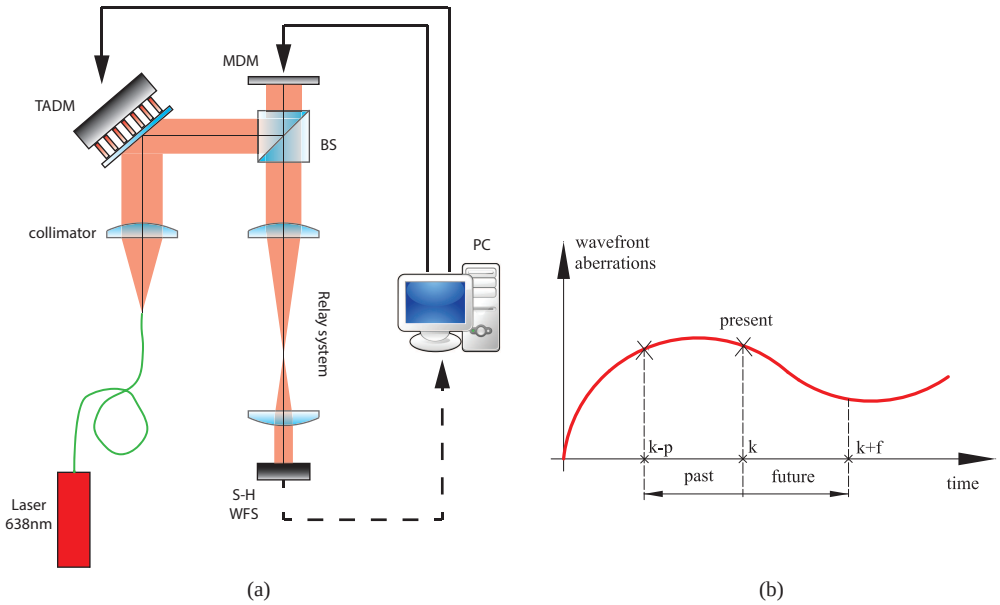
To demonstrate the proof of concept of predictive wavefront correction, we built an experimental AO system that is illustrated in Fig. 1(a). It consists of two DMs. The first DM is a commercial Membrane DM (MDM) produced by Adaptica Srl [25]. The MDM has 48 actuators and it was used as a wavefront correction mirror. This mirror corresponds to an active optical element in a lithographic machine that is used for wavefront correction. The second mirror is a prototype of a Thermally Actuated DM (TADM) that is developed by Eindhoven University of Technology [18, 19]. The TADM consists of the two cylindrical plates that are connected by 19 actuators. The actuators consist of aluminum rods with heating coils warped around them. When voltage is applied to an actuator it heats up and elongates. As it elongates, it exerts a mechanical force that deforms the mirror and a supporting back plate (for more details see [18]). The TADM was used to simulate the effect of TIWA in a lithographic machine. Accordingly, its inputs (voltages applied to actuators) serve as exposure conditions that determine wavefront aberrations.

The light of a semiconductor laser, working at the wavelength of $\lambda = 638$ [nm], is first collimated and then reflected by the TADM. The reflected light goes through the Beam Splitter (BS) and is again reflected by the MDM. The wavefront, at the surface of the MDM, is measured by a Shack-Hartmann Wavefront Sensor (S-H WFS) (Thorlabs WFS S300-14AR, 1.3 Mpixel, $\lambda/50$ rms accuracy). The MDM and the S-H WFS optically conjugate through a relay system consisting of two spherical lenses.

A predictive controller is implemented on a standard Personal Computer (PC). The PC sends two signals. One signal, that is calculated by the predictive controller, is sent to the MDM. At the same time, another actuation signal is sent to the TADM to introduce dynamical wavefront aberrations in the system. The predictive controller and the controller for the TADM are implemented in MATLAB. The predictive controller cannot influence the signal that is sent to the

#194428 - $15.00 USD Received 22 Jul 2013; accepted 19 Aug 2013; published 5 Sep 2013

TADM. The PC receives measurements from the S-H WFS. However, only at certain discrete time instants, the predictive controller is able to receive wavefront measurements. In order to indicate this, the line connecting the WFS and PC is dashed. This way, we simulate the situation in a real lithographic machine, where the wavefront aberrations can be measured only before and after the exposure process. All the connections in the system are established using LabVIEW environment.



(a) (b)

Fig. 1. (a) Experimental setup consisting of two DMs that was used to demonstrate the performance of the predictive control algorithm; (b) Past, present and future.

In the AO system, the control sampling period is $T = 2$s. This value is chosen on the basis of the step response analysis of the TADM (previously characterized in [26]). This sampling period divides the time t into k discrete time instants, that is, $t = kT$. The measured wavefront at the discrete time instant k is represented using a Zernike polynomial expansion (Noll [27]):

$$\Phi(x,y,k) = \sum_{i=1}^{36} \alpha_i(k) Z_i(x,y)$$

where $\Phi(x,y,k)$ is the spatial distribution of the wavefront phase, $\alpha_i(k)$ is the i^{th} coefficient of Zernike polynomial and $Z_i(x,y)$ is the i^{th} Zernike polynomial. In [26], it has been experimentally demonstrated that TADM can produce variation of essentially the first 9 Zernike coefficients (omitting the piston). We therefore focus on correction of these 9 Zernike coefficients which will be represented by a vector:

$$\mathbf{y}_T(k) = \begin{bmatrix} \alpha_2(k) & \dots & \alpha_{10}(k) \end{bmatrix}^T$$

The state-space model of the TADM has been identified in [26] and it has the following form:

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + Bv \tag{1}$$

$$\mathbf{y}_T(k) = C\mathbf{x}(k) \tag{2}$$

where $\mathbf{x}(k) \in \mathbb{R}^{10}$ is a state vector, $v \in \mathbb{R}$ is an input, and $A \in \mathbb{R}^{10\times10}$, $B \in \mathbb{R}^{10\times1}$ and $C \in \mathbb{R}^{9\times10}$ are the identified system matrices. In [26] it has been demonstrated that the state-space model

#194428 - $15.00 USD Received 22 Jul 2013; accepted 19 Aug 2013; published 5 Sep 2013

of the 4^{th} order (the dimension of the state $\mathbf{x}(k)$) is able to relatively accurately predict the dynamical behavior of the TADM. However, in order to achieve even better prediction accuracy in this paper we are using the identified model of the 10^{th} order.
The input of the TADM is $v = r_3^2$, where r_3 is a voltage applied to the 3rd actuator of the TADM (for more details consult [26]). For simplicity, we actuate only one channel of the TADM (the generalization for all 19 actuators of the TADM is straightforward).

From the prediction point of view, the state-space model of the TADM that is described by Eq. (1) and Eq. (2), represents the model of TIWA in a lithographic machine. Conceptually, the input v of this state-space model represents the exposure conditions that are inputs of the TIWA model (another modeling option is to see v as an intensity distribution on optical elements, however, for brevity we will not develop this idea in this paper). This input will be called *the disturbance input*. We assumed that the disturbance input is time independent because in a real system exposure conditions do not change during the exposure of a relatively large number of wafers. Because during exposure the light source is turned on and off with high temporal frequency, the intensity distribution on optical elements oscillates in time. However, it can be easily shown that the dynamical response of the thermoelastic system of PDEs to a high frequency intensity distribution can be approximated by response to a static intensity distribution. That is, from the modeling point of view high frequency intensity distribution can be approximated by a static intensity distribution.
Physically speaking, the state vector $\mathbf{x}(k)$ corresponds to the temperature and deformation of the optical elements in a lithography machine. That is, it corresponds to the states of the TIWA model.

The model of the MDM has the following form [28, 29]:

$$\mathbf{y}_M(k) = M\mathbf{u}(k), \qquad \mathbf{u}(k) = \begin{bmatrix} q_1^2(k) & q_2^2(k) & \dots & q_{48}^2(k) \end{bmatrix}^T \tag{3}$$

where $q_i(k) \in \mathbb{R}$ is a voltage applied to the i^{th} actuator of the MDM, $\mathbf{y}_M(k) \in \mathbb{R}^9$ is a shape of the MDM described by the 9 Zernike coefficients and M is the influence matrix. The vector $\mathbf{u}(k) \in \mathbb{R}^{48}$ will be called *the control input*. We identified M using the identification procedure explained in [30, 31]. The total wavefront $\mathbf{y}(k)$, that is measured by the S-H WFS, is a sum of the wavefronts produced by the two DMs:

$$\mathbf{y}(k) = \mathbf{y}_T(k) + \mathbf{y}_M(k) \tag{4}$$

In reality the total wavefront $\mathbf{y}(k)$ is corrupted by the S-H WFS measurement noise. For simplicity, in Eq. (4) we neglect the effect of the S-H WFS measurement noise. Combining Eq. (1), Eq. (3) and Eq. (4), we arrive at the model of the experimental setup:

$$\mathbf{x}(k+1) = A\mathbf{x}(k) + Bv \tag{5}$$

$$\mathbf{y}(k) = C\mathbf{x}(k) + M\mathbf{u}(k) \tag{6}$$

In order to define the prediction problem, we need to define the past and the future discrete-time horizons. At the discrete time instant k, *the past horizon* is the set of discrete-time instants $\{k-p, k-p+1, \dots, k-1, k\}$ and *the future horizon* is the set $\{k+1, k+2, \dots, k+f\}$, for some appropriately chosen values of p and f. The past and the future horizons are illustrated in Fig. 1(b). Next we define the following vectors:

$$\mathbf{v}_p = \underbrace{\begin{bmatrix} v \\ v \\ \vdots \\ v \end{bmatrix}}_{p \text{ entries of } v}, \quad \mathbf{u}_p = \begin{bmatrix} \mathbf{u}(k-p) \\ \mathbf{u}(k-p+1) \\ \vdots \\ \mathbf{u}(k) \end{bmatrix}, \quad \mathbf{y}_p = \begin{bmatrix} \mathbf{y}(k-p) \\ \mathbf{y}(k-p+1) \\ \vdots \\ \mathbf{y}(k) \end{bmatrix} \tag{7}$$

#194428 - $15.00 USD Received 22 Jul 2013; accepted 19 Aug 2013; published 5 Sep 2013

The vector $\mathbf{v}_p \in \mathbb{R}^p$ will be called *the vector of past disturbance inputs*. Similarly, $\mathbf{u}_p \in \mathbb{R}^{48(p+1)}$ and $\mathbf{y}_p \in \mathbb{R}^{9(p+1)}$ will be called *the vector of past control inputs* and *the vector of past wavefront measurements*, respectively. Next we define the following vectors:

$$\mathbf{v}_f = \underbrace{\begin{bmatrix} v \\ v \\ \vdots \\ v \end{bmatrix}}_{f \text{ entries of } v}, \mathbf{u}_f = \begin{bmatrix} \mathbf{u}(k+1) \\ \mathbf{u}(k+2) \\ \vdots \\ \mathbf{u}(k+f) \end{bmatrix}, \ \mathbf{y}_f = \begin{bmatrix} \mathbf{y}(k+1) \\ \mathbf{y}(k+2) \\ \vdots \\ \mathbf{y}(k+f) \end{bmatrix} \tag{8}$$

The vectors $\mathbf{v}_f \in \mathbb{R}^f$, $\mathbf{u}_f \in \mathbb{R}^{48f}$ and $\mathbf{y}_f \in \mathbb{R}^{9f}$ will be called *the vector of future disturbance inputs*, *the vector of future control inputs* and *the vector of future wavefronts*, respectively. We assumed that the same disturbance input v is an element of both $\mathbf{v}_p$ and $\mathbf{v}_f$ because in a real lithographic system the exposure conditions do not change for several hours.
The first predictive control problem is formulated as follows:

Predictive control problem 1. Using the past wavefront measurements, past control inputs, past and future disturbance inputs, and the model given by Eq. (5) and Eq. (6), at the time instant k determine the vector of future control inputs that will correct future wavefront aberrations.

In the second predictive control problem we will be interested in correction of wavefront aberrations when the disturbance input is not known a priori. This condition corresponds to the scenario when exposure conditions in a lithographic machine are either not modeled or they are not known a priori.

Predictive control problem 2. Using the past wavefront measurements, past control inputs, and the model given by Eq. (5) and Eq. (6), at the time instant k estimate the disturbance input v and determine the future control inputs that will correct future wavefront aberrations.

3. Predictive control strategies

In this section we solve predictive control problems 1 and 2. The solution of these problems is developed by deriving the multiple-steps ahead, AutoRegressive Exogenous inputs (ARX) predictor from Eq. (5) and Eq. (6) [24]. Similar predictive control strategy can be found in [32]. From Eq. (5) and Eq. (6) we have:

$$\underline{O}_p \mathbf{x}(k-p) = \mathbf{y}_p - \underline{I}_p \mathbf{v}_p - \underline{D}_p \mathbf{u}_p \tag{9}$$

where $\mathbf{y}_p$, $\mathbf{v}_p$ and $\mathbf{u}_p$ are defined in Eq. (7) and

$$\underline{O}_p = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^p \end{bmatrix}, \ \underline{I}_p = \begin{bmatrix} 0 & 0 & \dots & \dots \\ CB & 0 & \dots & \dots \\ CAB & CB & 0 & \dots \\ \vdots & & \ddots & \\ CA^{p-1}B & CA^{p-2}B & \dots & CB \end{bmatrix}, \ \underline{D}_p = \underbrace{\begin{bmatrix} M & 0 & 0 & \dots & 0 \\ 0 & M & 0 & \dots & 0 \\ 0 & 0 & M & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & M \end{bmatrix}}_{p+1 \text{ blocks}} \tag{10}$$

where $\underline{O}_p \in \mathbb{R}^{9(p+1)\times 10}$, $\underline{I}_p \in \mathbb{R}^{9(p+1)\times p}$ and $\underline{D}_p \in \mathbb{R}^{9(p+1)\times 48(p+1)}$. In Eq. (10), the matrices A, B, C and M are the system matrices of the model defined in Eq. (5) and Eq. (6). In order to

#194428 - $15.00 USD Received 22 Jul 2013; accepted 19 Aug 2013; published 5 Sep 2013

ensure that Eq. (9) can be uniquely solved for $\mathbf{x}(k-p)$, the length p of the past horizon has to be larger than the observability index [33] of the system defined by Eq. (5) and Eq. (6). That is, p has to be large enough such that $\underline{O}_p$ has full column rank. Assuming that this is the case, from Eq. (9) we have:

$$\mathbf{x}(k-p) = \underline{O}_p^{\dagger}\left(\mathbf{y}_p - \underline{I}_p\mathbf{v}_p - \underline{D}_p\mathbf{u}_p\right) \tag{11}$$

where $\underline{O}_p^{\dagger}$ denotes the pseudo-inverse of $\underline{O}_p$. From Eq. (1) we have:

$$\mathbf{x}(k) = A^p\mathbf{x}(k-p) + \underline{R}_{p-1}\mathbf{v}_p \tag{12}$$

where $\underline{R}_{p-1} = \begin{bmatrix} A^{p-1}B & A^{p-2}B & \dots & AB & B \end{bmatrix}$. Substituting Eq. (11) in Eq. (12), we obtain:

$$\mathbf{x}(k) = A^p\underline{O}_p^{\dagger}\mathbf{y}_p - A^p\underline{O}_p^{\dagger}\underline{D}_p\mathbf{u}_p + \left(\underline{R}_{p-1} - A^p\underline{O}_p^{\dagger}\underline{I}_p\right)\mathbf{v}_p \tag{13}$$

Similarly to Eq. (9), we have:

$$\mathbf{y}_f = \underline{O}_{f-1}A\mathbf{x}(k) + \overline{\underline{I}}_f\mathbf{v}_f + \underline{D}_{f-1}\mathbf{u}_f \tag{14}$$

where $\mathbf{y}_f$, $\mathbf{v}_f$ and $\mathbf{u}_f$ are defined in Eq. (8) and

$$\underline{O}_{f-1} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{f-1} \end{bmatrix}, \overline{\underline{I}}_f = \begin{bmatrix} CB & 0 & \dots & \dots \\ CAB & CB & 0 & \dots \\ \vdots & & \ddots & \ddots \\ CA^{f-1}B & CA^{f-2}B & \dots & CB \end{bmatrix}, \underline{D}_{f-1} = \underbrace{\begin{bmatrix} M & 0 & 0 & \dots & 0 \\ 0 & M & 0 & \dots & 0 \\ 0 & 0 & M & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & M \end{bmatrix}}_{f \text{ blocks}}$$

Substituting Eq. (13) in Eq. (14), we obtain:

$$\mathbf{y}_f = \underbrace{\underline{O}_{f-1}A^{p+1}\underline{O}_p^{\dagger}\mathbf{y}_p - \underline{O}_{f-1}A^{p+1}\underline{O}_p^{\dagger}\underline{D}_p\mathbf{u}_p - \underline{O}_{f-1}A^{p+1}\underline{O}_p^{\dagger}\underline{I}_p\mathbf{v}_p}_{\text{The effect of the initial state } \mathbf{x}(k-p)} + \underline{O}_{f-1}A\underline{R}_{p-1}\mathbf{v}_p + \overline{\underline{I}}_f\mathbf{v}_f + \underline{D}_{f-1}\mathbf{u}_f \tag{15}$$

The equation (15) tells us how the future wavefronts depend on the past wavefronts, past and future disturbance inputs and past and future control inputs. This equation will be called *the prediction equation*. The first three therms in the prediction equation originate from the initial state $\mathbf{x}(k-p)$ in Eq. (5). That is, in the general case the prediction cannot be done only on the basis of the model but the past wavefront measurements have to be taken into account. In the special case when $\mathbf{x}(k-p) = 0$ (that is, when initially the system is in the thermal equilibrium), the prediction equation can be simplified by neglecting the first three terms. In this special case, we do not need any past wavefront measurements to do the prediction.

The observability index of the identified state-space model of the TADM is 2. That is, in theory we only need 2 measurements samples of the past wavefront to predict the future behavior. However, in practice the past measurements are corrupted by the S-H WFS measurement noise. The negative effect of the measurement noise on the prediction performance, can be minimized by selecting larger p. For convenience we will write the prediction equation as follows:

$$\mathbf{y}_f = \mathbf{s} + \underline{D}_{f-1}\mathbf{u}_f \tag{16}$$

where

$$\mathbf{s} = \underline{O}_{f-1}A^{p+1}\underline{O}_p^{\dagger}\mathbf{y}_p - \underline{O}_{f-1}A^{p+1}\underline{O}_p^{\dagger}\underline{D}_p\mathbf{u}_p + \underline{O}_{f-1}A\left(\underline{R}_{p-1} - A^p\underline{O}_p^{\dagger}\underline{I}_p\right)\mathbf{v}_p + \overline{\underline{I}}_f\mathbf{v}_f \tag{17}$$

#194428 - $15.00 USD Received 22 Jul 2013; accepted 19 Aug 2013; published 5 Sep 2013

3.1. Solution of the prediction problem 1.

The first prediction problem will be solved using two approaches. In the first approach the future control inputs are determined by solving the following unconstrained least-squares problem:

$$\min_{\mathbf{u}_f}\{\mathbf{y}_f^T\mathbf{y}_f + \mathbf{u}_f^T W \mathbf{u}_f\} \tag{18}$$

where $\mathbf{y}_f$ is given in Eq. (16) and the weighting matrix $W \in \mathbb{R}^{48f\times 48f}$ is defined as follows:

$$W = \gamma Q Q^T, \quad Q = \begin{bmatrix} I & -I & 0 & \dots & & \\ 0 & I & -I & 0 & \dots & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & \dots & 0 & I & -I \\ & & & \dots & 0 & I \end{bmatrix} \tag{19}$$

In (19), $I \in \mathbf{R}^{48\times 48}$ is an identity matrix and γ is a positive regularization parameter. The entries of the matrix W penalize the difference between two consecutive elements of $\mathbf{u}_f$. This way, we can control the convergence speed of the predictive control algorithm. The weighting matrix can also take another form, that penalizes the energy of the future inputs or that improves the conditioning of the optimization problem, for details see for example [31]. The solution of (18) is:

$$\hat{\mathbf{u}}_f = -\left(\underline{D}_{f-1}^T \underline{D}_{f-1} + W\right)^{-1} \underline{D}_{f-1}^T \mathbf{s} \tag{20}$$

where $\mathbf{s}$ is defined in (17). The vector $\hat{\mathbf{u}}_f$ will be referred to as *the unconstrained predictive control input*. As it will be demonstrated experimentally (see Section 4), some of the elements of $\hat{\mathbf{u}}_f$ can be negative. Since the elements of $\hat{\mathbf{u}}_f$ are squares of the control voltages, negative entries of this vector are not feasible and we set them to zero. The consequence of this is a slower convergence of the predictive controller. In order to overcome this problem we add "hard" constraints to the cost function (18):

$$\begin{aligned} &\min_{\mathbf{u}_f}\{\mathbf{y}_f^T\mathbf{y}_f + \mathbf{u}_f^T W \mathbf{u}_f\} \\ &\text{subject to } \ \mathbf{a}_1 \preceq \mathbf{u}_f \preceq \mathbf{a}_2 \end{aligned} \tag{21}$$

where $\mathbf{a}_1 \in \mathbb{R}^{48f}$ is the vector of zeros, and $\mathbf{a}_2 \in \mathbb{R}^{48f}$ is a vector of ones. In Eq. (21), $\preceq$ denotes element-wise less-than-equal mathematical symbol. With the condition $\mathbf{u}_f \preceq \mathbf{a}_2$, we prevent saturation of the MDM. The solution of the optimization problem (21) will be referred to as *the constrained predictive control input*. Because the solution of this optimization problem cannot be found in the closed form, we have determined it using MATLAB function lsqlin().

3.2. Solution of the prediction problem 2.

In order to solve the second prediction problem, we need to estimate unknown disturbance input v from the past wavefront measurements. From Eq. (9) we have:

$$\mathbf{b} = L\mathbf{w}, \quad \mathbf{w} = \begin{bmatrix} \mathbf{x}(k-p) \\ v \end{bmatrix}$$

$$\text{where } \ \mathbf{b} = \mathbf{y}_p - \underline{D}_p \mathbf{u}_p, \quad L = \begin{bmatrix} \underline{O}_p & \underline{I}_p \mathbf{q} \end{bmatrix}, \quad \mathbf{q} = \underbrace{\begin{bmatrix} 1 & 1 & \dots & 1 \end{bmatrix}^T}_{\text{p entries}}$$

#194428 - $15.00 USD Received 22 Jul 2013; accepted 19 Aug 2013; published 5 Sep 2013

The unknown vector $\mathbf{w}$ consists of the initial state $\mathbf{x}(k-p)$ and the disturbance input. In reality the entries of the vector $\mathbf{y}_p$ are corrupted by a measurement noise. That is, the vector $\mathbf{b}$ is only approximately equal to $L\mathbf{w}$. We want determine $\mathbf{w}$ such that the difference $\mathbf{b}-L\mathbf{w}$ is as small as possible. This can be done by solving the following least-squares problem:

$$\min_{\mathbf{w}}\{(\mathbf{b}-L\mathbf{w})^T(\mathbf{b}-L\mathbf{w})\} \tag{22}$$

Assuming that L has full column rank, the solution of (22) is: $\hat{\mathbf{w}} = L^{\dagger}\mathbf{b}$. By estimating $\mathbf{w}$ we have at the same time estimated the initial state and the disturbance input v. By substituting the estimate of v in Eq. (18) or Eq. (21), and by solving these optimization problems we determine the future control inputs.

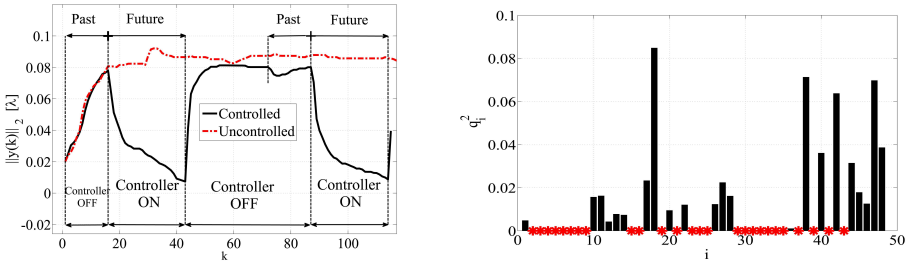
4. Experimental results

In this section, we present the experimental results of validating the predictive control strategy on the AO setup described in Section 2. To create nonzero initial states in the system, and thus to make the prediction problem more challenging, before each experiment the disturbance input of 15% of the maximal voltage value is applied for 30 [s]. In order to clearly distinguish the controlled and uncontrolled wavefronts, we do not control the MDM during the past horizons. That is, in the prediction equation (15) the past control input $\mathbf{u}_p$ is zero.

In Fig. 2(a) we show the performance of the unconstrained predictive controller. The performance of the predictive controller was quantified by computing the 2-norm of the measured Zernike coefficients (denoted by $\|\mathbf{y}(k)\|_2$ in all the figures). We assume that the disturbance input is known a priori. Two experiments were performed. In the both experiments, the disturbance input equal to the 35% of the maximal voltage value was applied to the TADM. In the first experiment, the MDM was not controlled. The uncontrolled wavefront is represented by a red dashed-dotted line in Fig 2(a). The wavefront $\|\mathbf{y}(k)\|_2$ starts from a value of 0.02 λ. This is because of the non-zero initial state in the system that was created by applying a 15% disturbance input before the beginning of the experiment. In about 20 discrete time samples (or 40s in total), the wavefront approximately reaches its steady state $\|\mathbf{y}(\infty)\|_2 \approx 0.09\lambda$

In the second experiment, the unconstrained predictive control input was calculated using Eq. (20) and it was applied to the MDM. This was done at two time instants: $k = 19$ and $k = 86$. The controlled wavefront is represented by a thick black line in Fig. 2(a). The future and past horizons were: $p = 17$ and $f = 30$. That is, 17 past measurements of the wavefront were taken, and the wavefront is predicted and controlled for 30 time instants in the future.

The elements of the unconstrained predictive control input, at an arbitrary time instant, are illustrated in Fig. 2(b). Some of the calculated voltages were negative, and they were set to zero (marked by stars in Fig. 2(b)).



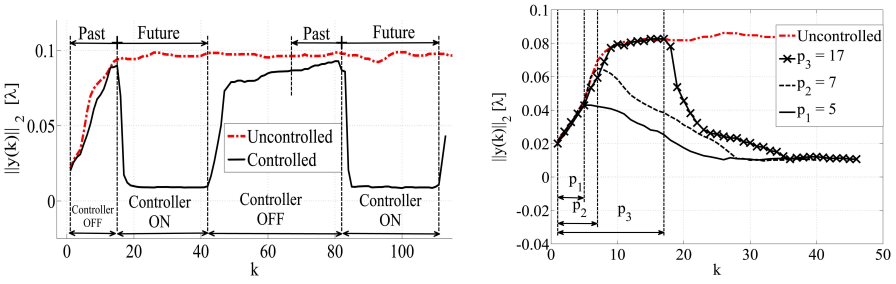
(a) (b)

Fig. 2. (a) Performance of the unconstrained predictive controller; (b) Calculated control input. Negative values of the control channels are set to zero.

#194428 - $15.00 USD Received 22 Jul 2013; accepted 19 Aug 2013; published 5 Sep 2013

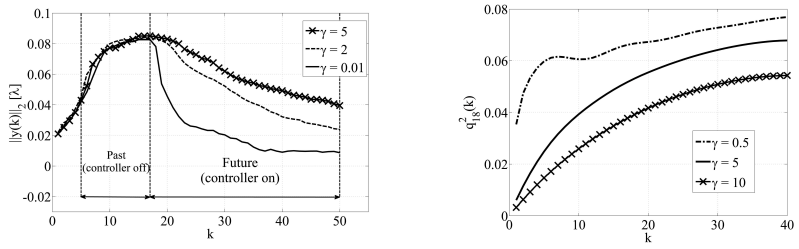
In Fig. 3(a) we show the performance of the constrained predictive controller. The predictive controller was calculated by solving the optimization problem (21). Similarly to the test of unconstrained predictive control strategy, two experiments were performed (the future and past horizons were identical to the ones used in the unconstrained case). From Figs. 2(a) and 3(a), it can be observed that both controllers are able to accurately correct wavefront aberrations (in both cases the converged value of $\|\mathbf{y}(k)\|_2$ is below 0.015λ). Furthermore, by comparing Figs. 2(a) and 3(a), it can be observed that the constrained predictive controller corrects wavefront aberrations faster than the unconstrained one. The unconstrained controller is slower because some of its voltages were set to zero. This way, we sacrificed the part of the performance of the unconstrained controller to get physically realizable voltages.

In Fig. 3(b), we show the performance of the unconstrained predictive controller for several values of p (the value of v is known a priori). Three past horizons were used $p_1 = 5$, $p_2 = 7$ and $p_3 = 17$. For smaller past horizons, the convergence rate is slower, and as the past horizon increases, the convergence is faster. This is because for larger p, we have more measurement data and consequently, we can better predict the future wavefront. In Fig. 4(a) we show how the convergence rate of the unconstrained predictive controller depends on the weighting parameter γ. In Fig. 4(b), we show how the convergence rate of an arbitrary voltage channel of the unconstrained predictive control input depends on γ. From Figs. 4(a) and 4(b), we see that when γ is increased, the convergence of the predictive controller gets slower. This is because larger γ implies stronger penalization of the difference between two consecutive future control inputs. Similar results were observed for the case of the constrained controller (for brevity we omit these results).



(a) (b)

Fig. 3. (a) Performance of the constrained predictive controller; (b) The performance of the unconstrained predictive controller for several values of past horizon p.



(a) (b)

Fig. 4. (a) Performance of the unconstrained predictive controller for different values of γ; (b) Convergence of an arbitrary channel of the unconstrained predictive control input.

Finally, in Fig. 5 we show the performance of the constrained predictive controller when unknown disturbance input is estimated (denoted by "estimated d. input"). For comparison we

#194428 - $15.00 USD Received 22 Jul 2013; accepted 19 Aug 2013; published 5 Sep 2013

also show the performance of the constrained predictive controller when the disturbance input is known a priori (denoted by ”real d. input”). The unknown disturbance input is estimated by solving the least-squares problem (22), and the estimate is substituted in the optimization problem (21) to derive the constrained predictive control input. Not surprisingly, the controller that is based on the estimated disturbance input has a (slightly) worse performance than the controller that is based on the ”real” disturbance input (difference of 0.01λ between the converged future wavefronts). This is because the estimate of the disturbance input is affected by the errors originating from the S-H WFS measurement noise and the model uncertainties of the TADM state-space model (described by Eq. (1) and Eq. (2)). These model uncertainties originate from the stochastic nature of the heat convection that occurs between the actuators of the TADM and the surrounding air [34]. We observed that the estimation quality and consequently the performance of the predictive strategy can be improved by increasing the length of the past horizon p. This is because more measurement data decrease the negative effect of the noise on the least-squares estimate of v.

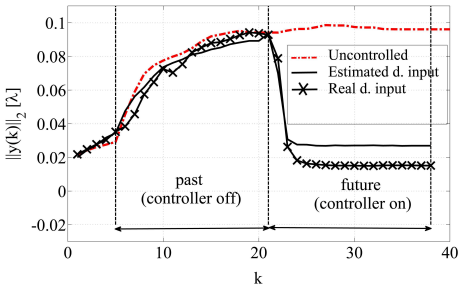


Fig. 5. The performance of the predictive controller when unknown disturbance input is estimated.

5. Conclusion

In this paper we have experimentally demonstrated a proof of concept for the predictive correction of TIWA in optical systems. The experimental results show that the predictive controller is able to correct wavefront aberrations using a relatively small number of past wavefront measurements. Furthermore, we have demonstrated that the predictive controller is able to correct wavefront aberrations even when the inputs of the wavefront prediction model are not known a priori. The proposed controller can be used in optical lithography as well as in other high-power optical systems where it is not possible to establish real-time feedback between a wavefront sensor and a controller.

Acknowledgment

This research is supported by the Dutch Ministry of the Economic Affairs and the Provinces of Noord-Brabant and Limburg in the frame of the ”Pieken in de Delta” program. We wish to thank R. Horsten and R. Pols for their technical support. Furthermore, we wish to thank S.K. Ravensbergen, P.C.J.N. Rossiele and M. Steinbuch from Eindhoven University of Technology for developing the TADM. Finally, we wish to thank A. van Dijsseldonk and J. van Schoot from ASML for providing us with useful information about the optical lithography.