Multiobjective optimization for target design in diffraction-based overlay metrology

YATING SHI,1 KUANGYI LI,1 XIUGUO CHEN,1,* PENG WANG,1 HONGGANG GU,1 Hao Jiang,1 Chuhanwei Zhang,1,2 and Shiyan Liu1,2

1State Key Lab for Digital Manufacturing Equipment and Technology, Huazhong University of Science and Technology, Wuhan 430074, China
2Wuhan Eoptics Technology Co. Ltd., Wuhan 430075, China
*Corresponding author: xiugochen@hust.edu.cn

Received 30 December 2019; revised 7 February 2020; accepted 13 February 2020; posted 14 February 2020 (Doc. ID 387066); published 18 March 2020

Overlay target design is an important issue in overlay metrology, whose aim is to probe the optimal overlay target to achieve good performance on measurement precision and accuracy even in the presence of process variation. In this paper, the target design problem is first formulated as a multiobjective optimization problem and then solved by the multiobjective genetic algorithm. The feasibility of the proposed method is verified based on simulations carried out on two overlay targets. The results reveal that measurements with high precision, accuracy, and process robustness could be achieved on the targets designed by the proposed method.

https://doi.org/10.1364/AO.387066

1. INTRODUCTION

The unceasing reduction of semiconductor device features derives increasingly stringent requirements on the overlay error budget [1–3]. The so-called overlay error refers to the misalignment of a feature from its intended position, which is aligned to an underlying patterned layer. As a rule of thumb, the overlay error should be less than about 20%–30% of the nominal critical dimension (CD) [3]. A large overlay error will result in a nonyielding device. According to the International Technology Roadmap for Semiconductors (https://www.itrs2.net), the requirement of the maximum allowable overlay error is 2 nm (3σ) at today’s feature size. The demand for tighter overlay control poses serious challenges to the conventional image-based overlay metrology technique. Over the past years, diffraction-based overlay (DBO) metrology techniques were developed to address the challenges [3–10].

The DBO techniques employed in overlay metrology, such as the normal incidence spectrophotometric reflectometry [4,5], the angle-resolved scatterometry [3,6], the Mueller matrix ellipsometry (MME) [7–12], and the phase-structured illumination technique [13], have already been used for the optical critical dimension (OCD) measurement in semiconductor manufacturing. However, the model-based data analysis methods developed for OCD metrology, such as the library search [14,15] and nonlinear regression approach [16,17], are not suitable for overlay metrology because solving an inverse problem is usually too time-consuming to meet the real-time requirement in overlay metrology. To this end, an empirical model-free approach that does not need to solve the inverse problem was widely adopted in overlay metrology [4,8–10]. The implementation of empirical DBO is performed on a specially designed target that consists of multiple cells with intentional overlay offsets. Each cell of the target is a stacked structure with two overlapping line gratings located on different patterned layers. The empirical approach relies on the linear response of the optical signal collected from a target with respect to the overlay offset within a small range. The signal conventionally used in overlay metrology is the reflectivity, such as the reflectance spectra of the 0th-order diffracted light [4,5] and the differential reflectance spectra of the ±1st-order diffracted light [3,6]. The Mueller matrix, which has already demonstrated advantages in asymmetric structure metrology [18,19], has recently been employed in overlay metrology [7–10,12] because the overlay target with nonzero overlay error is a typical asymmetric structure.

The implementation of empirical DBO is performed on a specially designed target to obtain measurement results with good performance. Instead of performing empirical trial-and-error experiments for multiple targets on wafer, simulation-based target design is especially important, as the best target could be selected at the pre-tapeout phase. Since the overlay target design is a typical optimization problem, the choice of optimization objectives (i.e., the metrics that reflect the quality of overlay metrology) is important. In the early research [4,6], the main metric concerned in overlay metrology is the measurement reproducibility (also referred to as precision) under the disturbance of random noise. Yang et al. [4] found that, when the intentional offset in an overlay target equals to a quarter of the grating pitch, better measurement precision could
be achieved. Ben-Dov et al. [20] and Tarshish-Shapir et al. [21] introduced two metrics used in overlay target design, i.e., the diffraction efficiency of the collected signal and the overlay error sensitivity, to improve the precision in overlay metrology. Except for the measurement precision, measurement robustness to process variation is another important metric in overlay metrology. The process variation refers to the deviation of the actual profile of a target from its design value, which is inevitable in the manufacturing process. The existence of process variation will contaminate the collected signals and thus introduce errors in overlay measurement results. At the advanced nodes, the measurement error caused by process variation may be too large and lead to unacceptable overlay measurement results. Therefore, the process robustness should also be considered in overlay target design to improve the immunity of measurement results to process variations [22].

In recent years, as the budget for overlay error continues to shrink, the measurement accuracy has been taken into consideration in overlay metrology [23,24]. However, Chen et al, [25] found that, if the measurement precision is emphasized too much, the accuracy will be compromised, which means that a better comprehensive performance in overlay metrology could be achieved through a trade-off between precision and accuracy. Since the behaviors of performance metrics in overlay metrology are inconsistent with each other, the overlay target design is a typical multiobjective optimization problem. The classical approach to solve a multiobjective optimization problem is to transform it into a single-objective optimization problem. One way to obtain the single objective function is to obtain the weighted summation of the multiobjective functions. In this case, the obtained solution is highly sensitive to the weight vector and demands that the user has thorough knowledge about the underlying problem. In another way, one of the objective functions is optimized by a single-objective optimization method; then, the obtained optimal results are further filtered step-by-step to meet the requirements of the other objective functions. However, solutions with the best comprehensive performance may not be obtained, as one of the objective functions (the one treated as the objective function in single-objective optimization problem) is emphasized too much in this approach.

In this paper, we resort to the multiobjective optimization genetic algorithm (MOGA) to realize the overlay target design, by which multiple objectives could be optimized simultaneously. The genetic algorithm (GA) is a metaheuristic inspired by the mechanism of natural selection and natural genetics [26]. Unlike a single-objective optimization problem, which has a definite solution, a set of points in solution space could be regarded as the solutions of a multiobjective optimization problem. Therefore, it seems natural to use GA in multiobjective optimization problems to capture a number of solutions simultaneously, since GA works with a population of points. In the GA, a group of randomly selected candidate solutions of the optimization problem are evolved toward better solutions through successive iterations. The candidates in each iterative step are termed as a “generation.” A new generation is formed from the previous generation through the bio-inspired operators, such as mutation, crossover, and selection. After several generations, the algorithm is expected to converge to the optimum or suboptimal solutions of the problem. Therefore, GA is an ideal candidate for solving multiobjective optimization problems [27,28] and is thus chosen to realize the target design in overlay metrology. It is worth noting that the collected signals from an overlay target depend not only on the geometric profile of the target but also on the measurement configuration (such as the wavelength, incident angle and/or azimuthal angle, etc.). In other words, the effects of target profile and measurement configuration on the measured signal are inseparable, which means that optimization of the target profile and measurement configuration should be taken into consideration in the target design process.

The remainder of this paper is organized as follows. In Section 2, the principle of empirical DBO is first revisited. Then, the overlay target design problem is formulated as a multiobjective optimization problem, and the overlay target design by MOGA is briefly introduced with an emphasis on important concepts of MOGA. In Section 3, simulations are performed on two prototypic overlay targets to verify the effectiveness of the proposed method. The results reveal the feasibility of MOGA in handling the overlay target design problem. Conclusions are drawn in Section 4.

2. METHOD

A. DBO Techniques in Overlay Metrology

The successful application of DBO techniques depends on the assumption that the collected signal \( I \) responds linearly to the overlay offset \( \delta \) in a small range and is zero when the overlay offset is absent, if the overlay offset is the only asymmetric feature in the structure under test. The signal \( I \) could be the summation of Mueller matrix elements from the off-diagonal \( 2 \times 2 \) blocks, such as \( M_{13} + M_{31} \) [7,10] that we used in this paper, or the difference between the reflectivity of the ±1st diffraction orders [3,6], etc. Based on this assumption, the relation between \( I \) and \( \delta \) could be expressed as

\[
I = K\delta,
\]

where overlay offset \( \delta \) denotes the total displacement between patterns in different layers, and the coefficient \( K \) represents the sensitivity of the signal \( I \) with respect to the overlay offset \( \delta \).

Figure 1 illustrates the layout and cross-section diagrams of an overlay target that is commonly used in DBO techniques. The overlay target is a grating-over-grating structure that consists of four cells. The line gratings in the top and bottom layer are of equal pitch \( \Lambda \), and \( r_1 \) and \( r_2 \) denote the line-to-pitch ratios of the top and bottom line gratings, respectively.  \( \varepsilon_x \) and \( \varepsilon_y \) denote the actual overlay error along the \( X \) and \( Y \) directions, respectively. In each of the cells, the top grating is intentionally shifted with respect to the bottom grating. Two cells with intentional offset \( D \) in opposite directions are used to measure the overlay error in the direction perpendicular to the lines of gratings. Without loss of generality, \( \varepsilon_x \) and \( \varepsilon_y \) are denoted as \( \varepsilon \) for simplicity. Therefore, the total overlay offset between the top and bottom gratings in two cells are \( \delta^+ = D + \varepsilon \) and \( \delta^- = D + \varepsilon \), respectively. By substituting \( \delta^+ \) and \( \delta^- \) into Eq. (1), the overlay error \( \varepsilon \) could be expressed as
where $I^+ = K(\pm D + \varepsilon)$ is the signal collected from the cell with an intentional offset of $\pm D$.

### B. Overlay Target Design by MOGA

The principle of DBO techniques seems simple; however, the situation in actual measurement is complicated; further, a lot of factors that may introduce errors into measurement results have to be taken into consideration. First, random noise is inevitably included in the measured signal. It can be easily found in Eq. (1) that small overlay sensitivity $K$ will amplify the effect of noise and thus lead to poor measurement precision. Second, contrary to the linear assumption, the relation between the diffraction signal $I$ and overlay offset $\delta$ is nonlinear in general cases. If one wants to obtain accurate results by empirical DBO techniques, good linearity has to be guaranteed at least in the range $\delta \in [-D + \varepsilon, D + \varepsilon]$.

Fig. 2 illustrates two types of process variations that are typically presented in the overlay target: the symmetric process variation and the asymmetric process variation. These variations will contaminate the overlay sensitivity $K$ and degrade the linearity of the relation in Eq. (1). Furthermore, the asymmetric process variation, as its name implies, will introduce additional asymmetric features in the target and thus lead to nonzero signal $I$, even when the overlay offset is absent. Therefore, the contamination caused by process variation, if not minimized, will significantly degrade the overlay measurement results. Based on the above analysis, instead of Eq. (1), the actual relation between the signal $I$ and overlay offset $\delta$ should be expressed as

$$I(a, p, \delta) = K(a, p)\delta + \Delta I_1(a, p) + \Delta I_p(a, p) + \Delta I_N,$$

where $a$ denotes a vector consisting of target profile parameters and measurement configuration; $p$ represents the process variation, such as $\Delta h$ and $\Delta d$ in Figs. 2(b) and 2(c); $\Delta I_1$ denotes the effect of imperfect linearity; $\Delta I_p$ represents the signal induced by asymmetric process variation; and $\Delta I_N$ is the random noise, which depends on the measurement instrument setup. The measurement configuration could be a combination of wavelength $\lambda$, incident angle $\theta$, and azimuthal angle $\phi$, as shown in Fig. 3.

By substituting Eq. (3) into Eq. (2), the measured overlay error $\varepsilon_m$ could be expressed as

$$\varepsilon_m(a, p, D) = \frac{I^+(a, p, D + \varepsilon_0) + I^-(a, p, -D + \varepsilon_0)}{I^+(a, p, D + \varepsilon_0) - I^-(a, p, -D + \varepsilon_0)} D,$$

where $\varepsilon_0$ denotes the actual overlay error. In order to obtain overlay measurement results with high precision, accuracy, and process robustness, the vector $a$ should be optimized to make the difference between $\varepsilon_m$ and $\varepsilon_0$, i.e., the measurement error $\varepsilon_m(a, p, D) = \varepsilon_m(a, p, D) - \varepsilon_0$, as small as possible and stable under the disturbance of random noise and process variations. It is worth noting that the intentional offset $D$ is usually fixed at a certain value [21]. As illustrated in our previous work [12], a positive correlation exists between the measurement error $\varepsilon_m(a, p, D)$ and intentional offset $D$. Therefore, if $D$ is treated as a parameter to be optimized, the optimal value of $D$ may be close to zero to obtain the smallest $\varepsilon_m$, which will make the two cells with offset $\delta^\pm = \pm D + \varepsilon$ degrade into one cell ($\delta^+ = \delta^- = \varepsilon$) and thus make the empirical DBO technique fail.

In order to solve the overlay target design problem, the problem should be first expressed mathematically. For an overlay target with a certain process variation $p$, $\mu(a, p)$ and $\sigma(a, p)$ are used to denote the mean value and standard deviation of the measured overlay error under the disturbance of random noise, respectively. Typically, different process variations will
lead to different measurement results. To achieve precise and accurate measurement results, \( \mu(a, p) \) and \( \sigma(a, p) \) obtained even under the worst case should be as small as possible. In addition, \( \mu(a, p) \) and \( \sigma(a, p) \) should be stable for any \( p \in \sum_p \), where \( \sum_p \) denotes the domain of process variation, to achieve process robust measurement results. Therefore, the target design problem in overlay metrology could be expressed as a multiobjective minimization problem:

\[
\hat{a} = \arg \min_{a \in \sum_a} \{ \| \mu(a, p) \|_\infty, \| \sigma(a, p) \|_\infty, \sigma(\mu(a, p)), \sigma(\sigma(a, p)) \}.
\]

for \( \forall p \in \sum_p \),

(5)

where \( \sum_a \) denotes the domain of variable \( a \), and \( \| \|_\infty \) denote the infinite norm of a vector. Since the infinite norm represents the element with the maximal absolute value in a vector, the “min-max” relation implicated in the first and second objectives in Eq. (5) indicates that the optimal result is workable even in the worst case. In other words, accurate and precise results could be achieved by minimizing the first and second objectives. Similarly, process robust results could be achieved by minimizing the third and fourth objectives in Eq. (5), since \( \sigma(\mu(a, p)) \) and \( \sigma(\sigma(a, p)) \) indicate the stability of measurement accuracy and precision under the disturbance of process variation. For simplicity, in the following part, the four objectives in Eq. (5) are denoted by \( f_1, f_2, f_3, \) and \( f_4 \), respectively.

For a single-objective optimization problem, one could attempt to find the best solution, which is absolutely superior to all other alternatives. However, for a multiobjective optimization problem, such as the problem described in Eq. (5), it is almost impossible to find a solution that could achieve the optimal for all the objectives simultaneously because the behavior of objectives is not consistent with each other. Mathematically, for a minimization problem with \( R \) objectives \( \{ z_1 = f_1(x), z_2 = f_2(x), \ldots, z_R = f_R(x) \} \), if \( f_j(x) \leq f_j(y) \) for all indices \( i \in \{1, 2, \ldots, R \} \), and \( f_j(x) < f_j(y) \) for at least one index \( j \in \{1, 2, \ldots, R \} \), then the solution \( x \) is said to dominate the solution \( y \) [26,27]. A solution is said to be Pareto optimal or nondominated if it is not dominated by any other solution in the solution space. Any improvement in any objectives is not possible for a Pareto optimal solution without sacrificing at least one of the other objectives. Generally, there is more than one Pareto optimal solution for a multiobjective optimization problem, and the objective function values of the Pareto optimal solutions are termed as the Pareto front, as shown in Fig. 4.

Since GA works with a population of points in the variable domain, it is capable of capturing the Pareto optimal solutions. As a widely used method, the nondominated sorting genetic algorithm (NSGA) [28,29] is an important implementation of MOGA, which adopts a mechanism called “Pareto ranking” to assign a proper fitness value to each candidate solution. As schematically shown in Fig. 4, the Pareto ranking procedure is illustrated as follows: assigning Rank 1 to the nondominated solutions and removing them from contention, then finding the next set of nondominated solutions and assigning Rank 2 to them and repeating this process until the entire population is ranked. For a minimization problem, a lower rank corresponds to a better solution. Many other Pareto ranking approaches have also been proposed over the past decades. One can consult [26,27] for more details about MOGA.

3. RESULTS AND DISCUSSION

In this paper, simulations are performed for three- and five-layer targets with overlay error \( e_0 \in [ -5, 5 \) nm. It should be noted that the heights of stacks in the overlay target are determined by the manufacturing process and thus are fixed at certain values. The intentional offset \( D \) is also fixed at a certain value, such as 30 nm in this paper. To evaluate the objectives in Eq. (5), random noise that obeys the normal distribution \( N(0, 0.001^2) \) is added on the simulated signal; then, 100 times repetitive measurements are performed to evaluate \( \mu(a, p) \) and \( \sigma(a, p) \). The signal \( I \) used in this paper is \( M_{13} + M_{31} \), which is calculated by in-house codes developed based on the rigorous coupled-wave analysis (RCWA) method [30–32]. The specific implementation of MOGA in this paper was carried out by using the “gamultiobj” function in MATLAB (version 2017a, The MathWorks, Inc., Natick, MA, USA), which is developed based on the NSGA-II [28,29]. The default number of iterative generations in gamultiobj is 100 times the number of variables to be optimized and here is set at 100. All the other parameters are set at their default values. The simulations are performed on a workstation equipped with double 2.0 GHz Intel Xeon CPUs.

A. Design of a Three-layer Overlay Target with Asymmetric Process Variation

The target studied in this section is a three-layer structure, as shown in Fig. 5. The top layer structure is an HSQ-resist grating with a height of \( H_1 \) and line-to-pitch ratio of \( r_1 \). The second layer is SiO\(_2\) film with a height of \( H_2 \). The pattern in the third layer is an Si grating with SiO\(_2\) in the groove, and the height and line-to-pitch ratio are denoted by \( H_3 \) and \( r_2 \), respectively. We assume that the heights of stacks are \( H_1 = H_3 = 150 \) nm and \( H_2 = 50 \) nm. In order to display the effect of optimization intuitively, two parameters (the pitch of gratings \( \Lambda \) and the azimuthal angle \( \psi \)) are optimized in this case. The ranges of pitch and azimuthal angle are \( \Lambda \in [200, 500] \) nm and \( \psi \in [0, 90] ^\circ \), which means that the variable domain \( \sum_a \) is a
2D space. The line-to-pitch ratios, wavelength, and incident angle are fixed at \( r_1 = r_2 = 0.5 \), \( \lambda = 425 \text{ nm} \), and \( \theta = 65^\circ \), respectively. The process variation considered in this case is the sidewall asymmetry \( \Delta d \) of the bottom Si grating, as depicted in Fig. 2(c). The domain of process variation \( \Xi = \{-4, 4\} \text{nm} \), which is divided by a step of 1 nm in the objective’s calculation process. It could be easily found that \( \Delta d = -4 \) or \( 4 \text{ nm} \) corresponds to a sidewall angle of about \( 87^\circ \) for the target under study.

Since there are four objectives in the optimization problem under study, the Pareto front achieved by MOGA would be four-dimensional. Thus, it is impossible to display the whole Pareto front intuitively. Instead, we exemplified several 2D cross-sections of the achieved Pareto front, as shown in Fig. 6. It could be observed that the four objectives all converge to small values for all the Pareto optimal solutions. It is worth noting that the Pareto optimal solutions are the mathematical solutions of the problem under study; these solutions could further be chosen based on the problem-related knowledge or factors. In this case, the optimal solutions should be further filtered by some predefined tolerance for objectives. For example, in this case, the tolerances for four objectives in Eq. (5) are set as 1, 0.1, 0.05, and 0.05 nm, respectively. Solutions satisfying these tolerances are shown in Table 1. We can find that the optimal solutions for the target under study.

![Cross-section diagram of the DBO target used in Example 1.](image1)

**Fig. 5.** Cross-section diagram of the DBO target used in Example 1.

**Fig. 6.** 2D cross-sections of the Pareto front achieved by MOGA in Example 1.

### Table 1. Pareto Optimal Solutions and Corresponding Objective Function Values in Example 1

<table>
<thead>
<tr>
<th>No.</th>
<th>( \Lambda ) (nm)</th>
<th>( \psi ) (°)</th>
<th>( f_1 ) (nm)</th>
<th>( f_2 ) (nm)</th>
<th>( f_3 ) (nm)</th>
<th>( f_4 ) (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>397.37</td>
<td>62.27</td>
<td>0.7162</td>
<td>0.0301</td>
<td>0.0478</td>
<td>0.0003</td>
</tr>
<tr>
<td>2</td>
<td>395.28</td>
<td>59.89</td>
<td>0.0950</td>
<td>0.0500</td>
<td>0.0202</td>
<td>0.0011</td>
</tr>
<tr>
<td>3</td>
<td>408.46</td>
<td>65.99</td>
<td>0.4387</td>
<td>0.0655</td>
<td>0.0155</td>
<td>0.0029</td>
</tr>
<tr>
<td>4</td>
<td>410.09</td>
<td>66.68</td>
<td>0.3326</td>
<td>0.0933</td>
<td>0.0176</td>
<td>0.0049</td>
</tr>
<tr>
<td>5</td>
<td>396.37</td>
<td>59.92</td>
<td>0.1379</td>
<td>0.0518</td>
<td>0.0292</td>
<td>0.0004</td>
</tr>
<tr>
<td>6</td>
<td>409.01</td>
<td>66.36</td>
<td>0.3785</td>
<td>0.0795</td>
<td>0.0145</td>
<td>0.0041</td>
</tr>
<tr>
<td>7</td>
<td>397.09</td>
<td>61.77</td>
<td>0.5719</td>
<td>0.0316</td>
<td>0.0485</td>
<td>0.0005</td>
</tr>
<tr>
<td>8</td>
<td>397.51</td>
<td>61.67</td>
<td>0.5094</td>
<td>0.0317</td>
<td>0.0492</td>
<td>0.0003</td>
</tr>
<tr>
<td>9</td>
<td>394.83</td>
<td>59.89</td>
<td>0.1541</td>
<td>0.0448</td>
<td>0.0325</td>
<td>0.0006</td>
</tr>
<tr>
<td>10</td>
<td>410.06</td>
<td>66.74</td>
<td>0.3157</td>
<td>0.0983</td>
<td>0.0152</td>
<td>0.0053</td>
</tr>
<tr>
<td>11</td>
<td>398.09</td>
<td>62.40</td>
<td>0.6879</td>
<td>0.0301</td>
<td>0.0440</td>
<td>0.0003</td>
</tr>
<tr>
<td>12</td>
<td>399.95</td>
<td>62.66</td>
<td>0.6215</td>
<td>0.0308</td>
<td>0.0356</td>
<td>0.0004</td>
</tr>
</tbody>
</table>

process robust overlay measurement results could be achieved on the target designed by MOGA.

To verify the effectiveness of the proposed method, relations between the signal \( M_{13} + M_{31} \) and overlay offset \( \delta \) are calculated for the Pareto optimal solutions displayed in Table 1. Figure 8 illustrates the relations of three typical optimal solutions, i.e., the solutions Nos. 1, 2, and 3 in Table 1. In each subfigure, five curves with different process variation \( \Delta d \) are displayed. It can be observed that the linear assumption between \( M_{13} + M_{31} \) and \( \delta \) is approximately achieved in the range \([-D + \varepsilon_0, D + \varepsilon_0]\), and the collected signal is basically immune to the process variation for these optimal solutions. Furthermore, the features shown in the figures are closely related to the value of the objective functions. For example, the linearity of curves in Fig. 8(b) is better than that in Fig. 8(a), which indicates better measurement accuracy. Accordingly, the objective \( f_1 \) of Solution No. 2 is better than that of Solution No. 1. The slope (i.e., the overlay sensitivity \( K \)) in Fig. 8(a) is higher than that in Fig. 8(b), which indicates better measurement precision. Accordingly, the objective \( f_2 \) of Solution No. 1 is smaller than
Fig. 7. Distribution of common logarithms of the objectives (a) $\log(f_1)$, (b) $\log(f_2)$, (c) $\log(f_3)$, and (d) $\log(f_4)$ in the domain $\Xi$. Data points marked with red circles correspond to the Pareto optimal solutions displayed in Table 1.

Fig. 8. Relations between the signal $M_{13} + M_{31}$ and overlay offset $\delta$ for Pareto optimal solutions (a) No. 1, (b) No. 2, and (c) No. 3 in Table 1.

that of Solution No. 2. The effect of process variation (which is denoted by objectives $f_1$ and $f_4$ together) in Fig. 8(c) is more significant than that in Figs. 8(a) and 8(b). We can find from Table 1 that, although the $f_3$ of Solution No. 3 is only one-third of the $f_3$ of Solution No. 1, the $f_4$ of Solution No. 3 is almost 10 times larger than that of Solution No. 1. A similar law could also be observed between solutions No. 2 and No. 3.

We also performed simulations for all the Pareto optimal solutions in Table 1 to mimic the actual measurement. For each solution, a random process variation $\Delta d$ and overlay error $\epsilon_0$ are selected in their own domains, and the corresponding Mueller matrices of two cells are calculated by RCWA. Then, signals $I^+$ and $I^-$ are obtained by adding random noise to the Mueller matrices collected from two cells of the target. Finally, the measured overlay error $\epsilon_m$ is solved by Eq. (4). The simulations are performed for 100 times, and the correlations between the measured overlay error $\epsilon_m$ and actual overlay error $\epsilon_0$ for Solution No. 1 are illustrated in Fig. 9 as an example. Obviously, an excellent linear correlation between $\epsilon_m$ and $\epsilon_0$ can be observed. The linear fitting results between $\epsilon_m$ and $\epsilon_0$ for all the optimal solutions are listed in Table 2. We can find that the coefficients of determination $R^2$ for all the solutions are close to one, and slopes and intercepts in fitted linear equations are almost close to one and zero, respectively. These results indicate that the overlay error $\epsilon_m$ measured on targets designed by the proposed method is close to the real value $\epsilon_0$ even under the disturbance of random noise and process variation.

B. Design of a Five-layer Overlay Target with Both Symmetric and Asymmetric Process Variation

The target studied in this case is a five-layer structure, as shown in Fig. 10. The top layer structure is the HSQ-resist grating with a height of $H_t$ and line-to-pitch ratio of $r_1$. The second to fourth layer are $Al_2O_3$, $Si_3N_4$, and $SiO_2$ films with a height...
of \( H_1, H_2, \) and \( H_3 \), respectively. The pattern in the fifth layer is an Si grating with SiO\(_2\) in the groove, and the height and line-to-pitch ratio are denoted by \( H_5 \) and \( r_2 \), respectively. We assume that the heights of stacks are \( H_1 = H_6 = 150 \) nm and \( H_1 = H_2 = H_3 = 50 \) nm. In this case, there are six parameters to be optimized, including the pitch \( \Delta \) of gratings, the line-pitch ratios \( r_1 \) and \( r_2 \) of the top and bottom gratings, the wavelength \( \lambda \), the incident angle \( \theta \), and the azimuthal angle \( \psi \), which means that the variable domain \( \mathbb{X} \) is a 6D space. The range of these parameters are \( \Delta \in [200, 500] \) nm, \( r_1 \in [0.3, 0.7] \), \( r_2 \in [0.3, 0.7] \), \( \lambda \in [400, 800] \) nm, \( \theta \in [45, 70] \)°, and \( \psi \in [20, 90] \)°. The process variation, including the sidewall asymmetry of bottom gratings \( \Delta d \) and the height of SiO\(_2\) film \( \Delta h_i \), is \( \Delta d \in [-4, 4] \) nm and \( \Delta h_i \in [-3, 3] \) nm. When calculating the objectives in Eq. (5), the domains of \( \Delta d \) and \( \Delta h_i \) are both divided by a step of 1 nm.

Figure 11 presents the 2D cross-sections of the Pareto front achieved by MOGA in Example 2.

![2D cross-sections of the Pareto front achieved by MOGA in Example 2](image)

The data point marked with the red circle is Solution No. 1. For each image, two variables float in their own domains, and the other four variables are fixed at the optimal values achieved by MOGA.

The selected solutions are displayed in Table 3. It could be observed that the optimal solutions in this case are located in the domain \( \Delta \in [436.01, 489.11] \) nm, \( f_1 \in [0.49, 0.51] \), \( f_2 \in [0.35, 0.53] \), \( \lambda \in [636.60, 717.53] \) nm, \( \theta \in [53.21, 59.35] \)°, and \( \psi \in [51.64, 83.22] \)°.

Table 3. Pareto Optimal Solutions and Corresponding Objective Function Values in Example 2

<table>
<thead>
<tr>
<th>No.</th>
<th>( \Delta ) (nm)</th>
<th>( r_1 )</th>
<th>( r_2 )</th>
<th>( \lambda ) (nm)</th>
<th>( \theta ) (°)</th>
<th>( \psi ) (°)</th>
<th>( f_1 ) (nm)</th>
<th>( f_2 ) (nm)</th>
<th>( f_3 ) (nm)</th>
<th>( f_4 ) (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>446.10</td>
<td>0.49</td>
<td>0.53</td>
<td>717.53</td>
<td>55.47</td>
<td>51.64</td>
<td>0.1662</td>
<td>0.1992</td>
<td>0.0046</td>
<td>0.0055</td>
</tr>
<tr>
<td>2</td>
<td>436.01</td>
<td>0.52</td>
<td>0.49</td>
<td>664.88</td>
<td>59.35</td>
<td>58.55</td>
<td>0.1577</td>
<td>0.1703</td>
<td>0.0077</td>
<td>0.0086</td>
</tr>
<tr>
<td>3</td>
<td>489.11</td>
<td>0.52</td>
<td>0.40</td>
<td>636.60</td>
<td>53.68</td>
<td>83.22</td>
<td>0.2150</td>
<td>0.1353</td>
<td>0.0382</td>
<td>0.0139</td>
</tr>
<tr>
<td>4</td>
<td>487.88</td>
<td>0.50</td>
<td>0.36</td>
<td>657.30</td>
<td>53.21</td>
<td>78.45</td>
<td>0.6029</td>
<td>0.0767</td>
<td>0.0891</td>
<td>0.0065</td>
</tr>
<tr>
<td>5</td>
<td>447.94</td>
<td>0.50</td>
<td>0.47</td>
<td>698.07</td>
<td>55.12</td>
<td>61.52</td>
<td>0.5156</td>
<td>0.1637</td>
<td>0.0985</td>
<td>0.0073</td>
</tr>
<tr>
<td>6</td>
<td>487.71</td>
<td>0.51</td>
<td>0.38</td>
<td>656.94</td>
<td>54.03</td>
<td>76.71</td>
<td>0.3569</td>
<td>0.1129</td>
<td>0.0488</td>
<td>0.0111</td>
</tr>
<tr>
<td>7</td>
<td>488.53</td>
<td>0.50</td>
<td>0.35</td>
<td>658.35</td>
<td>53.37</td>
<td>80.02</td>
<td>0.9325</td>
<td>0.0563</td>
<td>0.0980</td>
<td>0.0017</td>
</tr>
<tr>
<td>8</td>
<td>487.71</td>
<td>0.51</td>
<td>0.38</td>
<td>660.38</td>
<td>54.16</td>
<td>76.49</td>
<td>0.4407</td>
<td>0.1041</td>
<td>0.0691</td>
<td>0.0092</td>
</tr>
</tbody>
</table>
Fig. 12. Distribution of objectives in the variable domain $\mathfrak{D}_3$. Data points marked with red circles correspond to the Pareto optimal solution No. 1 displayed in Table 3. The first to fourth columns are the distributions of $\lg(f_1)$, $\lg(f_2)$, $\lg(f_3)$, and $\lg(f_4)$, respectively.

Fig. 13. Comparison of the measured overlay errors $\varepsilon_m$ and the actual overlay errors $\varepsilon_0$ for the Pareto optimal solution No. 1 in Table 3.

Table 4. Linear Fitting Results of Measured Overlay Errors and Actual Overlay Errors in Example 2

<table>
<thead>
<tr>
<th>No.</th>
<th>Fitted Linear Equation</th>
<th>$R^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$\varepsilon_m = 0.9892\varepsilon_0 + 0.0282$</td>
<td>0.9955</td>
</tr>
<tr>
<td>2</td>
<td>$\varepsilon_m = 0.9806\varepsilon_0 + 0.0189$</td>
<td>0.9963</td>
</tr>
<tr>
<td>3</td>
<td>$\varepsilon_m = 0.9938\varepsilon_0 - 0.0050$</td>
<td>0.9972</td>
</tr>
<tr>
<td>4</td>
<td>$\varepsilon_m = 0.9292\varepsilon_0 + 0.0109$</td>
<td>0.9986</td>
</tr>
<tr>
<td>5</td>
<td>$\varepsilon_m = 0.9697\varepsilon_0 - 0.0444$</td>
<td>0.9915</td>
</tr>
<tr>
<td>6</td>
<td>$\varepsilon_m = 0.9697\varepsilon_0 - 0.0444$</td>
<td>0.9978</td>
</tr>
<tr>
<td>7</td>
<td>$\varepsilon_m = 0.8646\varepsilon_0 + 0.0045$</td>
<td>0.9989</td>
</tr>
<tr>
<td>8</td>
<td>$\varepsilon_m = 0.9491\varepsilon_0 + 0.0032$</td>
<td>0.9979</td>
</tr>
</tbody>
</table>

targets designed by the proposed method is very close to the real value $\varepsilon_0$ even under the disturbance of random noise and process variation.

4. CONCLUSION

In this paper, the target design problem in overlay metrology based on the DBO techniques was first formulated as a multiobjective optimization problem, and the MOGA was introduced to probe the optimal combination of the target profile and measurement configuration, with which good performance in measurement accuracy, precision, and process robustness could be achieved. The summation of Mueller matrix elements from off-diagonal blocks, $M_{13} + M_{31}$, was exemplified as the indicator of overlay error to realize the target design. Simulations are carried out for two overlay targets to verify the effectiveness of the proposed method. The results clearly demonstrated the feasibility of MOGA in the target design problem in overlay metrology. The linearity and sensitivity between the collected signal and overlay error, as well as the immunity of measurement results to process variation, are all dramatically improved after optimization. Excellent linear relations could be observed between the actual overlay error and the overlay error measured on the targets designed by the proposed method. For the studied examples, results with measurement precision better than 0.2 nm and accuracy higher than 1 nm could be achieved by performing optimization.

Although $M_{13} + M_{31}$ is used as the indicator of overlay error in this paper, the proposed method could be readily extended to DBO techniques with other indicators, such as the summation of other Mueller matrix elements from off-diagonal blocks, the differential reflectance spectra of the ±1st-order diffracted light, etc. If the differential reflectance spectra of the ±1st-order diffracted light is used as the indicator of overlay error, a constraint between the grating pitch and wavelength of the incident beam should be considered in the optimization.
to ensure the $\pm 1$st-order diffracted light could be collected. Moreover, the proposed method could also easily be extended by adding the desired objectives to achieve the desired optimization goals. Therefore, the proposed target design method based on MOGA is expected to provide a more general and practical means to design the target utilized in overlay metrology. It is also worth noting that the design object considered in this paper is the profile of the overlay target on wafer; in addition, though not discussed in this paper, the corresponding pattern on lithographic mask could be readily calculated by performing inverse lithography [33].

**Funding.** National Natural Science Foundation of China (51525502, 51727809, 51775217, 51805193); China Postdoctoral Science Foundation (2019M652633); National Major Science and Technology Projects of China (2017ZX02101006-004); Natural Science Foundation of Hubei Province of China (2018CF057, 2018CBF559).

**Disclosures.** The authors declare no conflicts of interest.

**REFERENCES**