

Journal of
Micro/Nanolithography,
MEMS, and MOEMS

Nanolithography.SPIEDigitalLibrary.org

Overlay error statistics for multiple-exposure patterning

Allen H. Gabor
Nelson M. Felix

Overlay error statistics for multiple-exposure patterning

Allen H. Gabor^{a,*} and Nelson M. Felix^b

^aIBM Research, Yorktown Heights, New York, United States

^bIBM Research, Albany, New York, United States

Abstract

Background: The mathematical equations that explain overlay error of multiple-exposure patterning schemes have not been fully described in the literature and some commonly accepted methods lead to inaccurate estimated and/or measured overlay error.

Aims: Develop the proper mathematical framework, using a first principles statistical approach, so that engineers using multiple-exposure patterning can determine the overlay impact and overlay controls needed. Alert patterning community that grouped overlay metrology of multiple-exposures undermeasures the true overlay error.

Approach: Use image placement error and population-based statistics to enable a mathematical framework to be established that predicts the actual overlay error for an overlaying pattern that minimizes overlay error back to a pattern that is patterned with multiple-exposure patterning.

Results: The overlay error between two patterns is usually less than the root sum square of the two overlay error values of the patterns individually measured to a common prior pattern. Overlay error for a pattern minimizing back to multiple-prior patterns increases quickly as systematic overlay error between the prior patterns increases.

Conclusions: Controlling systematic overlay error between patterns of a multipatterned layer is important for subsequent patterns that need to minimize overlay error back to the composite multipatterned layer. The ratio between the overlay error determined with metrology and true overlay can be calculated.

© The Authors. Published by SPIE under a Creative Commons Attribution 4.0 Unported License. Distribution or reproduction of this work in whole or in part requires full attribution of the original publication, including its DOI. [DOI: [10.1117/1.JMM.18.2.021202](https://doi.org/10.1117/1.JMM.18.2.021202)]

Keywords: overlay error; multiple exposure; image placement error; population-based statistics; pooled variance; process assumptions; overlay metrology.

Paper 18104SS received Aug. 24, 2018; accepted for publication Feb. 28, 2019; published online Apr. 4, 2019.

1 Introduction

Multiple-exposure patterning^{1,2} is now a main-stream method used in the manufacturing of integrated circuits.³ While this paper focuses on pitch split multiple-exposure patterning, it is certainly not the only technique for patterning pitches below the diffraction limit ($0.5\lambda/\text{NA}$). For example, sidewall image transfer and directed self-assembly are alternative methods for resolving pitches below the single-exposure diffraction limit.^{4,5} In general, for pitch split multiple-exposure patterning, a lithography-process-lithography-etch (LPLE) scheme for double patterning will be used.⁶ Lithography-etch-lithography-etch⁷ and lithography-freeze-lithography-etch⁸ are two examples of these type of processes. While LPLE schemes are for double patterning, the general process for splitting a layer into n exposures is $[(n-1) * \text{LP}]LE$, where n is the number of exposures. In other words, the lithography-process portion can be repeated if splitting the layer into even more than two exposure steps is required. The following clarifications will aid readers in reading this paper:

- Multiple-exposure patterning, multipatterned, or multiple exposure all refer to a pitch split $[(n-1) * \text{LP}]LE$ processes, where $n \geq 2$ in this paper. To be specific, this

paper describes the overlay implications of pitch split multiple-exposure patterning but does not describe implications of other multiple-exposure patterning techniques.

- In this paper “layer” is used to refer to a functional layer, in the build of a semiconductor chip.
- “Pattern” is synonymous with layer in single-exposure systems. However, pattern is often the better term to use when discussing the overlay between exposures of a multipatterned layer. In this paper, we usually use pattern when describing an overlay between two patterns regardless of whether the overlay is between two distinct functional layers or between patterns of a multipatterned layer. This is consistent with SEMI standard P18-92 which uses the term “overlaying pattern.”⁹

So that readers, who are not well versed in overlay and minimizing overlay error in a modern semiconductor fabrication facility, can read this paper without having to consult overlay experts and study the referred to papers we have included expansive background material in Sec. 7.

1.1 Consequences of Overlay Error on Dimensional Error of Multipatterned Composite Layers

An error in the size of the space between lines occurs because of line critical dimension (CD) error with simple single-layer exposures. (We arbitrarily use lines as the feature being

*Address all correspondence to Allen H. Gabor, E-mail: allen.gabor@ibm.com

patterned and space between as the dependent feature.) In other words, the CD determines not only the size of the features being measured but also the space between features. In addition, with single exposure, the CD variation of the lines is equal to the dimensional variation of the spaces between the lines.

Using multiple exposures to pattern a composite layer, results in the determination of space variation being more complicated. (Composite layer refers to a functional layer, in the build of a semiconductor chip, that is formed by multiple-exposure patterning.) It was recognized by the industry, early in the development of multiple-exposure patterning, that the overlay requirements for meeting the dimensional targets, of the indirectly patterned features of the composite layer, can be challenging to achieve when using multiple-exposure patterning.¹⁰⁻¹² As an example, for layers that are pitch split into two separate exposures, both the CDs of the features being patterned and the overlay error between the two exposures being used for the composite layer determine the final spacing between the features being patterned.¹³ Previous literature has focused on the effect of overlay error in multiple-exposure patterning on the dimension of the feature not being directly patterned. For example, both Arnold et al.¹⁰ and Hazelton et al.¹² have each described mathematically the effect of overlay error and CD error of the two exposures, of a double patterning process, on the indirectly patterned features.

In the case where the multiple-exposure patterning is setting the CD of the dielectric for a metal layer, if overlay error between the multiple exposures increases, the width-variation of the metal lines will also increase. If there is a systematic translation error between the two (or more) patterns of the composite layer, a bi- (or multi-) modal distribution of metal line-widths will result and will affect circuit timing. Alternatively, if the multiple-exposure patterning is setting the CD of the metal lines, a systematic translation error between the multiple-exposed patterns can result in a multimodal distribution of dielectric width between the metal lines and lead to dielectric breakdown. While a systematic translation error between the multiple exposures will result in a multimodal width distribution of the indirectly patterned features, random overlay error will simply increase the width variation of the indirectly patterned features (and not make the width distribution multimodal). Thus with multiple-exposure patterning, whether the overlay error is random or systematic, the dimensional error of the indirectly patterned feature will be larger than that of the directly patterned feature.

It should be noted that besides translation error, there are other spatially systematic overlay errors (magnification, rotation, trapezoid, etc.).¹⁴ This paper (and past literature) includes all systematic error besides translation error as being part of the 3-sigma variation. This simplification likely results in cases of incorrectly estimated overlay error variation. We encourage those in the community to research this area and publish further work on how systematics, other than translation error, affect overlay error especially with multiple-exposure patterning.

1.2 Notation Used for Overlay and Image Placement Terms

In this paper, several terms used in deriving the mathematical formulas that describe overlay error of multiple-exposed

systems look similar. This section defines and clarifies the differences between these terms and introduces the concept of effective 3-sigma in overlay error control.

\vec{ip}_1 is defined as the difference a feature designed on pattern 1 has between its position on a reference grid and the actual value on a wafer. (Note: we specify pattern 1, but 1 is for illustration only and can be any pattern or layer, e.g., 2, 3, ..., C1, C2, ..., etc.) The \vec{ip}_1 is commonly referred to as the image placement error of pattern 1, but it is also called registration.⁹ \vec{IP}_1 is the average image placement error of pattern 1 ACFWL (across chip, field, wafer, and lot/lots). Thus the case of the text in \vec{ip}_1 and \vec{IP}_1 indicates whether the image placement error is at a specific (X, Y) point or the average image placement error, respectively. σ_1 is defined as the standard deviation of all the image placement errors of pattern 1.

In the one-dimensional (1-D) representation of image placement [Fig. 1(a)], the average image placement error \vec{IP}_1 is represented by the vertical line that is placed at the image placement error with the highest count. In the two-dimensional (2-D) representation [Fig. 1(b)], \vec{IP}_1 is at the center of the circle and the circle represents all the possible x and y image placement combinations having the same probability, i.e., there is no correlation between x and y image placements. Specifically, all points along the 3-sigma circle are equally probable. If a 2-sigma circle were illustrated and examined, all points along that circle would also be equally probable but would have greater probability than points on the 3-sigma circle. This 2-D graphical representation will be further described in Sec. 4.1 to develop the fundamental mathematical equations that describe overlay error of an overlaying pattern to a prior pattern.

Overlay error between an overlaying pattern and prior pattern is a vector quantity designated by $\vec{ol}_{2 \rightarrow 1}$, where:

- The lower case “ol” designates that the overlay error is for a specific (X, Y) point on a wafer.
- “2” represents the overlaying pattern.
- “1” represents the prior pattern.

The standard deviation $\sigma_{2 \rightarrow 1}$ includes ACFWL overlay error variation. (Note: the notation for overlay standard

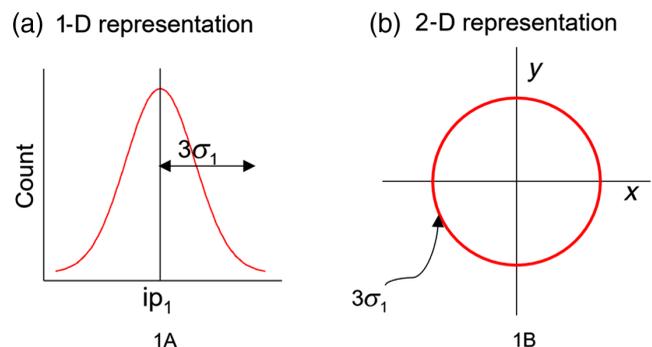


Fig. 1 (a) A one-dimensional representation of the image placement error in one orientation and (b) a two-dimensional representation of the image placement errors in both X and Y . \vec{ip}_1 is the image placement error and $3\sigma_1$ is three times the standard deviation of the image placement error.

deviation specifies both the overlaying pattern and the prior pattern, whereas the notation for image placement standard deviation specifies only one layer/pattern, thus enabling the reader to tell when overlay standard deviation is being used and when image placement standard deviation is being used.) $\overrightarrow{OL_{2 \rightarrow 1}}$ represents the systematic translation error and is a vector quantity. The OL of the vector quantity $\overrightarrow{OL_{2 \rightarrow 1}}$ is in capital letters to denote average overlay error rather than overlay error at a specific point. $OL_{2 \rightarrow 1}$ (without the arrow as it is a nonvector quantity) is used to estimate the effective 3-sigma overlay error that takes into account the effect of not having the systematic translation error equal to zero. To calculate the effective 3-sigma overlay error ($OL_{2 \rightarrow 1}$) the absolute value of the systematic translation error is added to three times the standard deviation of the overlay error [Eq. (1)]. This effective 3-sigma value can be used as the statistical process control (SPC) control limit, i.e., if the measured $OL_{2 \rightarrow 1}$ is greater than the control limit value, rework will be triggered. Examining Eq. (1), $\overrightarrow{OL_{2 \rightarrow 1}}$ is only a component of $OL_{2 \rightarrow 1}$ and $OL_{2 \rightarrow 1}$ can be a large value even when the systematic translation error is zero, i.e., the variation could be large but the systematic translation error equal to zero

$$OL_{2 \rightarrow 1} = |\overrightarrow{OL_{2 \rightarrow 1}}| + 3\sigma_{2 \rightarrow 1}. \quad (1)$$

Thus, it is important to note whether an arrow, representing that the overlay being referred to is a vector quantity, is present over the OL. If the arrow is above the OL, it is only one component of the effective 3-sigma overlay. Table 1 summarizes the notation used in this paper for overlay and image placement terms.

1.3 Overview of Paper

Because of the recognized importance of overlay in multiple-exposure patterning, many papers have delved in to the overlay effects of multiple-exposure patterning on CD control.¹⁰⁻¹³

This paper instead focuses on the impact of multiple-exposure patterning to:

1. a lithographer's capability to meet overlay specifications and
2. an overlay metrologist's ability to measure the overlay error.

To accomplish this, a mathematical framework is developed to not only estimate the overlay error within the composite layer (Sec. 3), but to also estimate the effect that the composite layer will have on subsequent layers that need to minimize overlay error to the composite layer (Sec. 4).

2 Interlayer and Intralayer Consequences of Multipatterned Layers on Overlay Error

This section describes the interlayer and intralayer overlay error consequences of choosing different overlay error minimization schemes of a multipatterned system using the classic assumption that the different overlay error components are random independent variables. It also sets the stage for what is fully developed in Sec. 4: minimizing the overlay error of an overlaying pattern to the union of all the prior patterns.

Interlayer effects with multiple-exposure patterning are caused by the fact that any layer minimizing overlay error back to a multipatterned layer has a prior pattern with multiple-image placement signatures to minimize back to. For example, if a contact layer was exposed with two exposures, a layer that needs to minimize overlay error back to contacts has two patterns to minimize back to. Minimizing back to only one of the prior exposures results in the overlay error of the overlaying pattern to the other prior pattern(s) being degraded.

Before continuing, it is important to discuss naming of the overlaying patterns and prior patterns. While in Sec. 1.2, we used SEMI standard⁹ numbers "2" and "1" to refer to the overlaying pattern and prior pattern, respectively, in the rest of this section, we will be examining a more complex system that has multiple-prior patterns. SEMI standard

Table 1 Notation used in this paper for overlay and image placement terms.

Term	Description	Vector quantity?
$\overrightarrow{ip_1}$	Image placement error at a specific (X, Y) point on a wafer for a pattern 1	Yes
$\overrightarrow{IP_1}$	Mean image placement error for a pattern 1 averaged ACFWL	Yes
σ_1	Standard deviation of image placement error for a pattern 1	No
$\sigma_{1A \cup 1B}$	Standard deviation of image placement error of a composite layer formed by the union of patterns 1A and 1B (see Sec. 4.3). Note 1A and 1B represent two patterns of a multiexposed composite layer	No
$\overrightarrow{ol_{2 \rightarrow 1}}$	Overlay error at a specific (X, Y) point on a wafer between an overlaying pattern (2) and a prior pattern (1)	Yes
$\overrightarrow{OL_{2 \rightarrow 1}}$	Systematic translation error between an overlaying pattern (2) and a prior pattern (1) averaged ACFWL	Yes
$\sigma_{2 \rightarrow 1}$	Standard deviation of overlay error between overlaying pattern (2) and a prior pattern (1) ACFWL	No
$OL_{2 \rightarrow 1}$	Effective 3-sigma overlay error between an overlaying pattern (2) and a prior pattern (1) that takes into account both the variation of overlay error and systematic translation error	No

terminology means that often a design layer is an overlaying pattern when it is initially patterned and then a prior pattern for subsequent exposed design layers. This makes handling multiexposed systems using SEMI standard terminology difficult. Because of this SEMI standard terminology issue, this paper will often use a pattern naming convention, illustrated in Fig. 2, when examining a multiexposed system. For all text and mathematical equations developed, we specify what is the overlaying pattern and what is the prior pattern(s). For example, when it is written “C2 to B overlay,” the first pattern written (C2) is the overlaying pattern and the second pattern written (B) is the prior pattern. Similarly, when “ $\overrightarrow{ol_{D1 \rightarrow C2}}$ ” is written, D1 is the overlaying pattern and C2 is the prior pattern. The two examples above illustrate that a pattern (C2) can be both an overlaying pattern and a prior pattern depending on what pattern combinations are being examined.

Figure 2(a) shows the case where the overlay error of an overlaying pattern (D1) is minimized to only one (C1) of the prior patterns. This minimization scheme results in the overlay error of the overlaying pattern to the other prior pattern (C2) being minimized indirectly through the minimization of three different mask pairs: D1 to C1, C1 to B, and C2 to B. Thus the D1 to C2 relationship is referred to as third order. At any given (X, Y) point on a wafer, the overlay error between D1 and C2 can be determined by

$$\overrightarrow{ol_{D1 \rightarrow C2}} = \overrightarrow{ol_{D1 \rightarrow C1}} + \overrightarrow{ol_{C1 \rightarrow B}} - \overrightarrow{ol_{C2 \rightarrow B}}. \quad (2)$$

The subtraction of $\overrightarrow{ol_{C2 \rightarrow B}}$ from the other terms in the determination of $\overrightarrow{ol_{D1 \rightarrow C2}}$ is done because overlay error is a vector quantity that changes direction (sign) if the “overlaying” pattern and “prior” patterns are reversed. To be specific, $\overrightarrow{ol_{2 \rightarrow 1}} = -\overrightarrow{ol_{1 \rightarrow 2}}$. If overlay errors are all independent random errors, the standard deviation of the overlay error between D1 and C2 can be estimated by taking the root sum square (RSS) of the standard deviations of all the error sources of the components shown in Eq. (2). (Note: while in this section, overlay error components are treated as independent random errors, Sec. 3 describes how the errors are often

correlated and further develops the mathematics for the correlated case.) The RSS of the standard deviations is shown in

$$\sigma_{D1 \rightarrow C2} = \sqrt{\sigma_{D1 \rightarrow C1}^2 + \sigma_{C1 \rightarrow B}^2 + \sigma_{C2 \rightarrow B}^2}. \quad (3)$$

It is important to go through this two-step process of Ref. 15:

- First, determining the equation that describes the error sources that combine to estimate the error of interest.
- Second, if it is a simple addition (or subtraction) of error sources and they are independent random variables, then the standard deviation of the error of interest can be estimated by taking the RSS of the standard deviations of the combining error sources.

Understanding how the overlay variation for multipatterned layers relates to single-layer overlay capability is desirable. If all the overlay error sources have standard deviations equal to that of single-layer overlay error standard deviation capability ($\sigma_{D1 \rightarrow C1} = \sigma_{C1 \rightarrow B1} = \sigma_{B1 \rightarrow C2} = \sigma_{SL2 \rightarrow SL1}$, where $\sigma_{SL2 \rightarrow SL1}$ is defined as the single-layer overlay error standard deviation capability) and are independent random variable, then $\sigma_{D1 \rightarrow C2} = \sqrt{3} \sigma_{SL2 \rightarrow SL1}$. The $\sqrt{3}$ times the single-layer overlay error standard deviation capability often prevents critical design rules from being supported with this third-order control scheme. If instead direct overlay minimization between C2 and C1, and second-order overlay minimization of C2 to B is used, then D1 to C2 overlay error becomes second order and the factor changes to $\sqrt{2} \sigma_{SL2 \rightarrow SL1}$ [Fig. 2(b)]. This $\sqrt{2}$ times the single-layer overlay error standard deviation capability still may be larger than the technology can allow. Because controlling the overlaying pattern to only one of the prior patterns leaves the overlay error to other prior patterns being indirectly minimized, many times the overlay error of an overlaying pattern is minimized to all prior patterns of a pitch split layer simultaneously [Fig. 2(c)].¹⁶ Minimizing the overlaying pattern to the union of the prior patterns is often an excellent strategy and the statistics of this scheme will be described in Sec. 4.

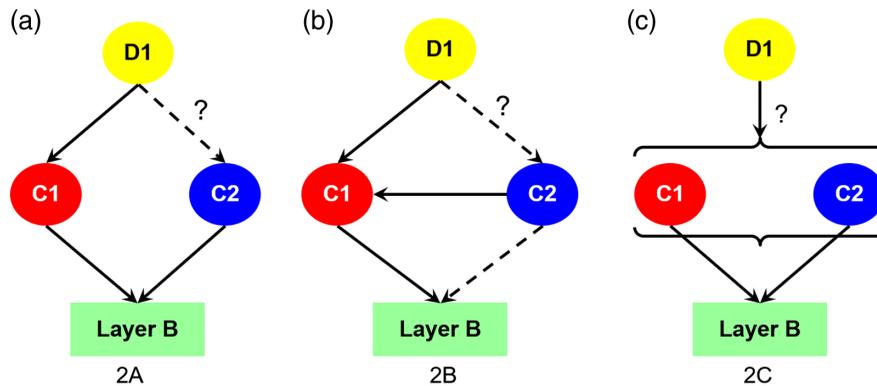


Fig. 2 Overlay error for an overlaying pattern D1 is minimized back to a prior pattern C that was exposed with two masks C1 and C2. Solid arrows are used to designate direct overlay minimization. Dashed arrows are used to designate indirect overlay minimization. Three different cases for D1 minimizing back to C are shown. (a) The case where D1 is third order to C2. (b) The case where D1 is second order to C2 at the expense of C2 being second order to B. (c) The preferred overlay minimization strategy where overlay error is minimized to the union of C1 and C2 enabling C2 to B to remain first order.

3 Estimating Overlay Error When Errors are not Random and/or Independent

This section describes what the correct mathematics are for relating the overlay errors that are directly measured and controlled and those that are indirectly controlled (and often not even measured) in the manufacturing of semiconductor chips using multiple-exposure patterning. It is also shown that the classic RSS approach (that was described in Sec. 2) is often an overestimate of the actual overlay error since the overlay error components are not usually independent random variables.

3.1 Estimating Indirectly Controlled Overlay Error When There is No Systematic Translation Error Between Directly Minimized Patterns

Section 2 reviewed how if one knows what error sources sum together, and all the component errors are independent random variables, one can RSS the standard deviations of error components together to estimate the standard deviation of the total error. Figure 3(a) shows the multiple-exposure patterning situation that will be analyzed in this section. Specifically, overlaying pattern C is split into two exposures C1 and C2, which both measure and minimize overlay error to a prior pattern B. Figure 3(a) shows a very common situation where C1 to B and C2 to B overlay error minimization is even more critical than C1 to C2 overlay. Specifically, the minimization to prior pattern B means that C1 to C2 overlay is indirect, and thus often not as low as it could be if direct overlay minimization between C1 and C2 was used. As described in Sec. 2, it is common practice to RSS the standard deviations $\sigma_{C1 \rightarrow B}$ and $\sigma_{C2 \rightarrow B}$ to estimate the overlay error standard deviation of C1 to C2 ($\sigma_{C1 \rightarrow C2}$). At any given (X, Y) point in an exposure field $ol_{C2 \rightarrow C1} = ol_{C2 \rightarrow B} - ol_{C1 \rightarrow B}$, so the error components satisfy at least parts of each of the two rules outlined in Sec. 2 for when RSS of error components can be used. However, Fig. 3(b) shows that the measured 3-sigma overlay between C1 and C2 is significantly lower than the C1 to C2 3-sigma overlay error estimated by the RSS of the C1 to B and C2 to B measured 3-sigma overlay error. The reason for this discrepancy is in order to RSS the standard deviations of the two error

components, the error sources not only have to sum together but also be independent random variables.¹⁵ In the case of C1 to B and C2 to B, the overlay errors are not necessarily independent variables. For example, if overlaying pattern C1 has a component of overlay error to prior pattern B, caused by B, overlaying pattern C2 will also have that error because B is common to both. Such a case could result if the mask used for prior pattern B has an image placement residual vector field signature that leads to both overlay error between C1 and B and overlay error between C2 and B. However, the degraded overlay error caused by the image placement residual vector field signature of prior pattern B does not impact the overlay error of C1 to C2. [As will be shown in Sec. 4.1 and Eq. (7), the fundamental definition of overlay is the difference in image placement between the two patterns, and therefore, the image placement error of B does not impact the overlay error of C1 to C2.] Another example is if C1 and C2 were exposed on the same scanner. In this case, they will both have an error component that the common scanner implements. This common error can come from the fact that C1 and C2 were exposed with common aberrations¹⁷ (assuming the illumination for C1 and C2 are similar), wafer chucks, reticle holders, baseline offsets, alignment offsets, pattern polarity, etc. In such situations, one cannot simply RSS the error sources because they are correlated and one must determine the correlation factor ρ . The determination of the correlation factor is beyond the scope of this paper but is an area we encourage the community to do research on and publish best practices. Once the correlation factor is determined, Eq. (4) can be used to estimate overlay error between C1 and C2. [Ref. 15 includes the case where the simple summing of the variances has to also include the correlation factor]

$$\sigma_{C2 \rightarrow C1} = \sqrt{(\sigma_{C1 \rightarrow B})^2 + (\sigma_{C2 \rightarrow B})^2 - 2\rho\sigma_{C1 \rightarrow B}\sigma_{C2 \rightarrow B}} \quad (4)$$

Equation (4) can be used to calculate an overlay error standard deviation between C1 and C2 of zero when the overlay error between C1 and B and C2 and B are perfectly correlated ($\rho = 1$) and equal in magnitude ($\sigma_{C1 \rightarrow B} = \sigma_{C2 \rightarrow B}$). Thus in an extreme case even if the overlay error between C1 and B and the overlay error between C2 and B are large

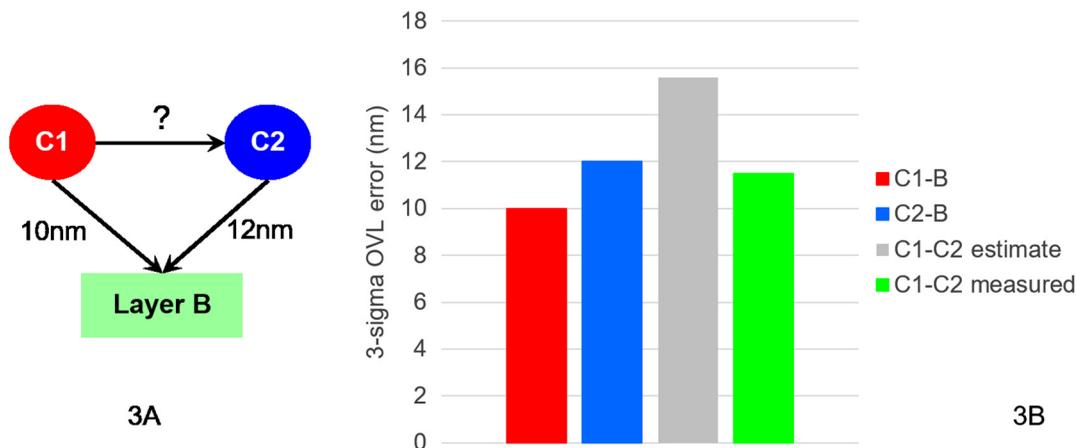


Fig. 3 (a) The case where a overlaying pattern C is exposed with two masks each minimizing overlay error to prior pattern B. C1 to C2 overlay error is “second order” but a RSS of the standard deviations of C1 to B and C2 to B overestimates the overlay error between C1 and C2 as shown in (b).

values, the overlay error between C2 and C1 will be low if $\sigma_{C1 \rightarrow B}$ and $\sigma_{C2 \rightarrow B}$ are highly correlated and equal in value.

3.2 Estimating Indirectly Controlled Overlay Error When There is Systematic Overlay Error Between Directly Minimized Patterns

Section 3.1 reviewed the mathematics that should be used to estimate indirectly controlled overlay error when there is no systematic translation error between the measured patterns. The effect of systematic translation error is addressed now. In the example shown in Fig. 3, if C1 to B and C2 to B not only had a standard deviations of $\sigma_{C1 \rightarrow B}$ and $\sigma_{C2 \rightarrow B}$, respectively, but also a systematic translation errors of $\overrightarrow{OL_{C1 \rightarrow B}}$ and $\overrightarrow{OL_{C2 \rightarrow B}}$ then the systematic translation error between C2 and C1 becomes:

$$\overrightarrow{OL_{C2 \rightarrow C1}} = \overrightarrow{OL_{C2 \rightarrow B}} - \overrightarrow{OL_{C1 \rightarrow B}}. \quad (5)$$

Substituting C2 for 2 and C1 for 1 in Eq. (1) and using the values of sigma and systematic translation error calculated in Eqs. (4) and (5), respectively, the estimated effective 3-sigma overlay error becomes:

$$OL_{C2 \rightarrow C1} = |\overrightarrow{OL_{C2 \rightarrow B}} - \overrightarrow{OL_{C1 \rightarrow B}}| + 3\sqrt{(\sigma_{C1 \rightarrow B})^2 + (\sigma_{C2 \rightarrow B})^2 - 2\rho\sigma_{C1 \rightarrow B}\sigma_{C2 \rightarrow B}}. \quad (6)$$

4 Mathematics to Estimate Overlay Error of an Overlaying Pattern to Multiple-Prior Patterns

Section 3 reviewed the statistics of estimating overlay error between two patterns that both minimize back to the same prior pattern. This section investigates the proper statistics to use when an overlaying pattern minimizes overlay error to the union of multiple-exposed prior patterns. It is demonstrated that whether the overlay error between multiple-exposed prior patterns is random or systematic influences the capability of subsequent overlaying patterns to minimize overlay error to the composite prior pattern.

As discussed in Sec. 2, one cannot start taking the RSS of the overlay error standard deviations of prior measured patterns, without confirming the error sources sum together. In the case of D1 minimizing overlay error to the union of C1 and C2 [Fig. 2(c)], there is no equation relating $\overrightarrow{ol_{D1 \rightarrow (C1UC2)}}$ to $\overrightarrow{ol_{C1 \rightarrow B}}$, $\overrightarrow{ol_{C2 \rightarrow B}}$, and/or $\overrightarrow{ol_{C1 \rightarrow C2}}$ overlay. Without such an equation, the RSS of overlay error component values has no basis. In the rest of Sec. 4, we will start from fundamental image placement error and then bring in the concept of population-based statistics to build an infrastructure for estimating overlay error of an overlaying pattern to the union of multipatterned prior patterns.

4.1 Overlay Error Determined from Image Placement Error for a Single-Layer Exposure

Fundamentally, the overlay between a feature of an overlaying pattern and a feature of a prior pattern is defined [Eq. (7)] as a difference in image placement between the two patterns⁹

$$\overrightarrow{ol_{2 \rightarrow 1}} = \overrightarrow{ip_2} - \overrightarrow{ip_1}. \quad (7)$$

Similarly, the mean overlay between all features of an overlaying pattern and prior pattern is expressed by

$$\overrightarrow{OL_{2 \rightarrow 1}} = \overrightarrow{IP_2} - \overrightarrow{IP_1}. \quad (8)$$

A broader discussion of overlay as defined by SEMI Standard P18-92 [Eq. (7)] is necessary. SEMI standard P18-92 is defined so the overlaying pattern and prior pattern can have a designed nonzero offset. Thus a nonzero difference in image placement values calculated with Eq. (7) does not necessarily mean there is an overlay error. For the remainder of this paper, however, it is assumed that:

1. The target value of overlay is zero, i.e., the designed value of $\overrightarrow{ip_2} = \overrightarrow{ip_1}$.
2. Any overlay value other than zero is an overlay error.

To be specific, although Eqs. (7) and (8) are defined as “overlay” (consistent with SEMI standard P18-92) any non-zero value is assumed to be an overlay error in this paper.

The relationship between image placement error and overlay error has been previously described.^{18,19} As Proglar et al. pointed out, image placement error has many sources including “lens aberration induced pattern shifts, reticle registration errors, and exposure tool placement variations via wafer and field systematic/random components.” The systematics can include “wafer/field translation, rotation, magnification, etc. errors”

Figure 4 shows a graphical representation of the overlay error between pattern 2 and pattern 1. The systematic translation error vector is illustrated in Fig. 4(a), which shows how the average field of an overlaying pattern 2 is displaced from a prior pattern 1. Note that for simplicity we have illustrated a systematic translation error that has a Y component only, i.e., the average X overlay error is zero. However, along with calculating the systematic translation error, determining the standard deviation of the overlay error ACFWL is key to

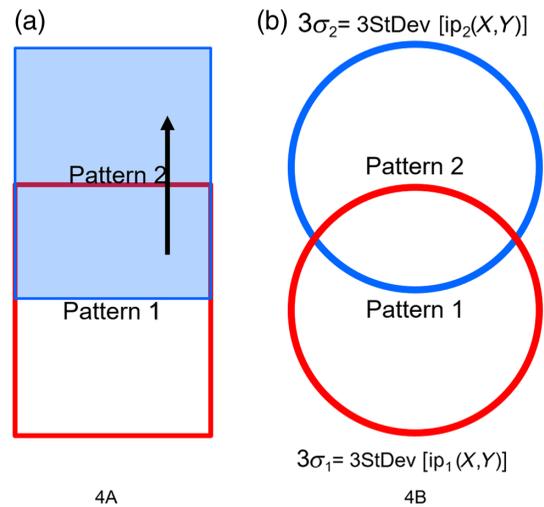


Fig. 4 (a) The exposure fields of two patterns 1 and 2 that have a translation error in the Y orientation and (b) while there is the average translation error shown in (a) there is also overlay error variation within each field due to the image placement variation of each pattern.

getting a complete estimate of overlay error. The variation in overlay error is illustrated in Fig. 4(b). The centers of the image placement error distribution for pattern 1 and pattern 2 are displaced from each other. The specific amount of the displacement of the centers is equal to the systematic translation error vector shown in Fig. 4(a). The diameter of the circles in Fig. 4(b) represents three times the standard deviation of all image placement error values for pattern 1 (red circle) and pattern 2 (blue circle).

A few comments on independent random variables and Fig. 4(b): independent and random means that an image placement error represented by a point at the top of the 3-sigma circle representing pattern 2 has equal probability of being paired to the pattern 1 image placement error represented at the top of the pattern 1, 3-sigma circle as that at the bottom (or any other point) of the pattern 1, 3-sigma circle. Importantly, any given pattern 2 image placement error has the greatest probability of being paired with the average pattern 1 image placement error represented by the center of the pattern 1 circle, assuming the distribution is Gaussian (as Gaussian distributions have more counts in the center of the distribution and continually have fewer counts as one looks further away from the average of the distribution). Also it is important to note that although the average overlay error in X is zero, that due to the variation in image placement error of pattern 2 and pattern 1, and the assumption that they are independent random variables, there is a distribution of overlay error vectors around the systematic translation error in both X and Y . Because the difference in image placement errors of the two patterns is equal to the overlay error between the two patterns [Eq. (7)], the two image placement standard deviations can be root sum squared to determine the standard deviation of the overlay error between pattern 2 and pattern 1

$$\sigma_{2 \rightarrow 1} = \sqrt{(\sigma_2)^2 + (\sigma_1)^2}, \quad (9)$$

$$\sigma_{2 \rightarrow 1} = \sqrt{2(\sigma_{ip})^2}. \quad (10)$$

Equation (10) can be used to calculate $\sigma_{2 \rightarrow 1}$ if all the sigmas are equal to σ_{ip} (where ip denotes image placement error). To calculate the 3-sigma effective overlay error ($OL_{2 \rightarrow 1}$), we add the systematic translation error determined with Eq. (8) to three times the sigma of overlay error determined with Eq. (10). Thus the effective 3-sigma overlay error is calculated with Eq. (11) following the same logic described at the end of Sec. 3.2

$$OL_{2 \rightarrow 1} = |\vec{IP}_2 - \vec{IP}_1| + 3 * \sqrt{2(\sigma_{ip})^2}. \quad (11)$$

4.2 Overlay of the D1 to C1 and C2 Union When Mean Image Placement Error for C1 is Same as C2

In general, if distribution C is composed of the union of distributions A and B , and A and B have the same average value and equal population, then the standard deviation of distribution C is given by

$$\sigma_C = \sigma_{A \cup B} = \sqrt{\frac{(\sigma_B)^2 + (\sigma_A)^2}{2}}. \quad (12)$$

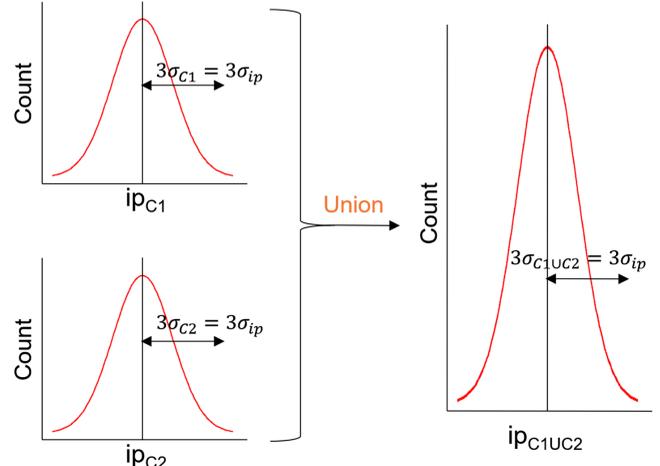


Fig. 5 When two distributions C1 and C2 combine, if they have the same means and standard deviations, the new distribution will have the same standard deviation.

Note that Eq. (12) is not an RSS, as Eq. (9) is, but rather a root mean square, which is the method to estimate the standard deviation of a combined population (when A and B have the same average value and equal counts). Figure 5 illustrates what occurs when two patterns C1 and C2 combine to form a new population, when $\vec{IP}_{C1} = \vec{IP}_{C2}$ and $\sigma_{C1} = \sigma_{C2} = \sigma_{ip}$. Specifically, the new combined population of image placements will have the same standard deviation and same average as those of the C1 and the C2 populations

$$\sigma_{C1 \cup C2} = \sqrt{\frac{(\sigma_{C1})^2 + (\sigma_{C2})^2}{2}} = \sqrt{\frac{(\sigma_{ip})^2 + (\sigma_{ip})^2}{2}} = \sigma_{ip}. \quad (13)$$

Following the steps outlined in Sec. 4.1 [for deriving Eqs. (9) and (10)], Eq. (14) can be derived and used to calculate the D1 to C1/C2 union [illustrated in Fig. 2(c)] overlay error standard deviation in the case where:

1. No mean image placement error exists between C1 and C2 [as required when using Eq. (12)].
2. The image placement errors of D1, C2, and C1 are equal.
3. It is understood that the prior pattern 1 in Eq. (9) can represent not only a single-exposed pattern, but also a composite pattern formed by multiple-exposure patterning:

$$\begin{aligned} \sigma_{D1 \rightarrow C1 \cup C2} &= \sqrt{(\sigma_{D1})^2 + \frac{(\sigma_{C1})^2 + (\sigma_{C2})^2}{2}} \\ &= \sqrt{2(\sigma_{ip})^2}. \end{aligned} \quad (14)$$

Note when the criteria established above is met (no mean error and equal standard deviations for image placement error), the overlay error of D1 to the C1/C2 union is equal to that of the single-layer overlay of D1 to C1 or the single-layer overlay of D1 to C2:

$$\begin{aligned}\sigma_{D1 \rightarrow C1} &= \sqrt{(\sigma_{D1})^2 + (\sigma_{C1})^2} = \sqrt{2(\sigma_{ip})^2} \\ &= \sqrt{(\sigma_{D1})^2 + (\sigma_{C2})^2} = \sigma_{D1 \rightarrow C2}.\end{aligned}\quad (15)$$

4.3 Overlay Error of D1 to the C1 and C2 Union When Mean Placement Error for C1 and C2 Differ

The overlay capability of D1 to a C1 and C2 union is degraded from single-layer overlay when the mean image placement error of C1 is different than the mean image placement error of C2. The field of population-based statistics has developed mathematics to estimate the pooled standard deviation of a new population when means or counts are not the same for two distributions that are combining. The general equation is given by Eq. (16),^{20,21} where N_A and N_B are the counts in each individual population (and are significantly greater than 10) and μ_A and μ_B are the means of the two populations:

$$\sigma_{AUB} = \sqrt{\frac{N_A(\sigma_A)^2 + N_B(\sigma_B)^2}{N_A + N_B} + \frac{N_A N_B}{(N_A + N_B)^2}(\mu_A - \mu_B)^2}.\quad (16)$$

Note the difference in the means term in Eq. (16). With this term, Eq. (16) enables the user to determine how much the standard deviation increases as the modes of the distribution separate from each other due to different mean values. Equation (16) can be utilized to estimate the standard deviation of image placement error of a C1/C2 union shown in Fig. 2(c). The case when population counts of C1 and C2 are equal ($N_{C1} = N_{C2}$) but mean IP error exists between C1 and C2 is illustrated (Fig. 6.). In such a case, using Eq. (16), the image placement standard deviation of the C1/C2 union is given by

$$\sigma_{C1UC2} = \sqrt{\frac{(\sigma_{C1})^2 + (\sigma_{C2})^2}{2} + \frac{1}{4}(\overline{IP}_{C1} - \overline{IP}_{C2})^2}.\quad (17)$$

The standard deviation of the overlay error between D1 and the union of C1/C2 is given by

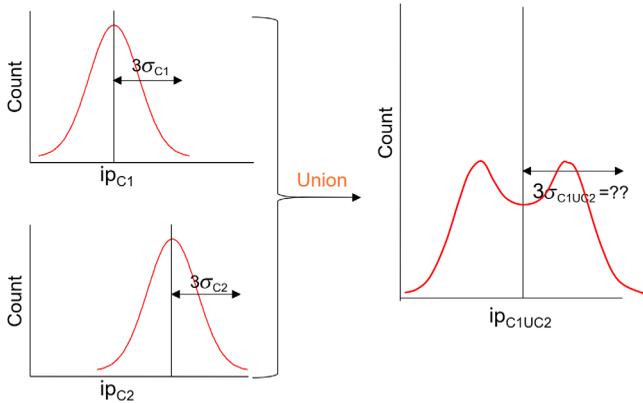


Fig. 6 When two distributions C1 and C2 combine, if they have the different means, counts, or standard deviations, the standard deviation of the combined distribution can be estimated utilizing Eq. (16).

$$\sigma_{D1 \rightarrow C1UC2} = \sqrt{(\sigma_{D1})^2 + \frac{(\sigma_{C1})^2 + (\sigma_{C2})^2}{2} + \frac{1}{4}(\overline{IP}_{C1} - \overline{IP}_{C2})^2}.\quad (18)$$

Equation (18) is derived using the image placement standard deviation of the C1/C2 union calculated with Eq. (17) and the fundamental definition of overlay error from the difference in image placement errors defined in Eq. (7). Then by applying the same methods used to estimate the effective 3-sigma overlay error in Sec. 3.2 [Eq. (6)], the effective 3-sigma overlay error between D1 and the union of C1/C2 is given by

$$\begin{aligned}OL_{D1 \rightarrow C1UC2} &= |\overline{IP}_{D1} - \overline{IP}_{C1UC2}| \\ &+ 3\sqrt{(\sigma_{D1})^2 + \frac{(\sigma_{C1})^2 + (\sigma_{C2})^2}{2} + \frac{1}{4}(\overline{IP}_{C1} - \overline{IP}_{C2})^2}.\end{aligned}\quad (19)$$

Equations (18) and (19) illustrate the importance of minimizing the systematic translation error ($\overline{IP}_{C1} - \overline{IP}_{C2}$) between prior patterns if the overlaying pattern (D1 in the case being illustrated) is to achieve good overlay capability. Figure 7 plots the effect of systematic translation error between prior patterns on the effective 3-sigma overlay error for the pattern that is to minimize back to a union of prior patterns as determined by Eq. (19) when the following two requirements are met:

1. The image placement standard deviations of D1, C1, and C2 are all equal to the single-layer overlay error standard deviation capability divided by $\sqrt{2}$, i.e., $\sigma_{D1} = \sigma_{C1} = \sigma_{C2} = \frac{\sigma_{SL2 \rightarrow SL1}}{\sqrt{2}}$. This results in both D1 to C1 and D1 to C2 overlay error standard deviations being equivalent to the single-layer overlay error standard deviation capability, i.e., $\sigma_{D1 \rightarrow C1} = \sigma_{D1 \rightarrow C2} = \sigma_{SL2 \rightarrow SL1}$.
2. $|\overline{IP}_{D1} - \overline{IP}_{C1UC2}|$ is equal to the single-layer overlay error capability $\sigma_{SL2 \rightarrow SL1}$.

Requirements 1 and 2 are summarized in

$$\begin{aligned}\sigma_{SL2 \rightarrow SL1} &= |\overline{IP}_{D1} - \overline{IP}_{C1UC2}| = \sqrt{(\sigma_{D1})^2 + (\sigma_{C1})^2} \\ &= \sqrt{(\sigma_{D1})^2 + (\sigma_{C2})^2}.\end{aligned}\quad (20)$$

When the requirements documented in Eq. (20) are met, the effective 3-sigma overlay error becomes $4\sigma_{SL2 \rightarrow SL1}$. While the mismatch between “3-sigma” and “ $4\sigma_{SL2 \rightarrow SL1}$ ” can appear to be an error, it is exactly what effective 3-sigma is intended to represent. Specifically, when there is a systematic translation error equal to $+\sigma_{SL2 \rightarrow SL1}$, the range that is needed to capture 99.7% of the points, (where the range is centered at the target overlay value, which is usually zero) will be from $-\sigma_{SL2 \rightarrow SL1}$ to $+\sigma_{SL2 \rightarrow SL1}$. Similarly, if the systematic translation error is $-\sigma_{SL2 \rightarrow SL1}$ away from its target the range that captures 99.7% of the points will move to $-\sigma_{SL2 \rightarrow SL1}$ to $+\sigma_{SL2 \rightarrow SL1}$.

By looking at the curves for the different single-layer overlay capabilities plotted in Fig. 7, one can see the effect

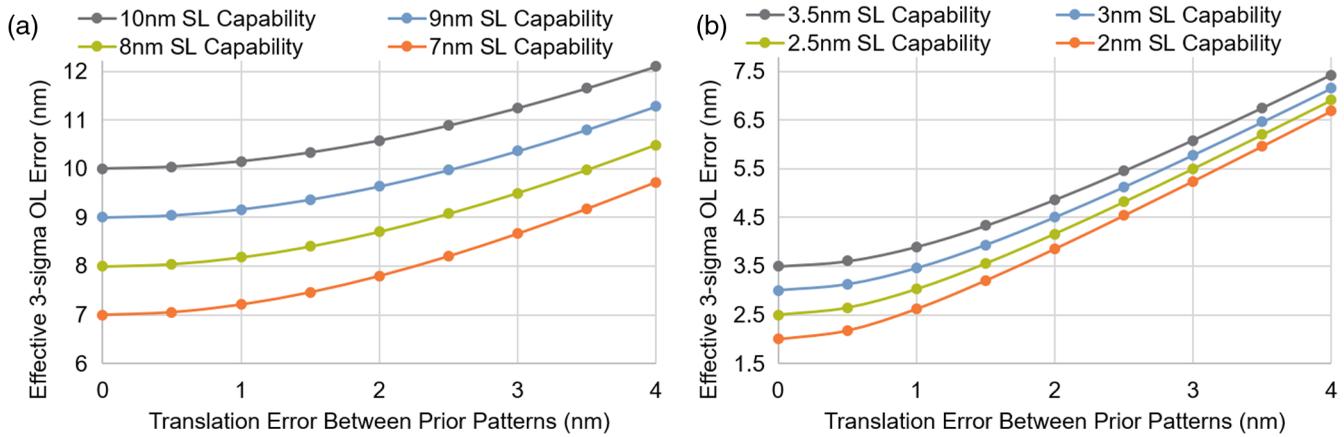


Fig. 7 The graphs are determined utilizing Eq. (19) and assume that: (a) all image placement standard deviation equals the single-layer overlay capability divided by the $\sqrt{2}$ and (b) $|\overrightarrow{IP}_{D1} - \overrightarrow{IP}_{C1/C2}|$ is equal to a quarter of the single-layer overlay capability.

that prior patterns, with intralayer systematic translation error ($\overrightarrow{IP}_{C1} \neq \overrightarrow{IP}_{C2}$), have on the overlay error of an overlaying pattern to the composite prior pattern. For a process that has an effective 3-sigma single-layer capability of 10 nm ($\sigma_{SL_2 \rightarrow SL_1} = 2.5$ nm), a 4-nm systematic translation error, within the union of the prior patterns, results in approximately a 2-nm increase in overlay error between an overlaying pattern and the union of prior patterns [Fig. 7(a)]. However, that same systematic translation error between the prior patterns increases the effective 3-sigma overlay by more than 4 nm for a process that has a base effective 3-sigma overlay error capability of 2 nm [Fig. 7(b)]. Thus as overlay specifications become smaller, controlling the systematic translation error between prior patterns becomes more important. To illustrate the use of these equations and the charts in Fig. 7 assume the following:

- Layers C1 and C2 from Fig. 2(c) represent a contact layer that is pitch split into two exposures.
- Layer B represents a gate layer.
- The overlay error of each of the two contact layers is minimized back to gate.
- C1 to gate is shipped with a -0.75 -nm systematic translation error and C2 to gate with a $+0.75$ -nm systematic translation error.

Using the above assumptions, one can calculate that the C1 to C2 overlay error has a systematic translation error of 1.5 nm. If the overlay process assumption (PA) for D1 to the C1/C2 union is 3.5 nm (effective 3-sigma), a process that has single-layer capability of 2.5 nm will be needed due the fact that C1 and C2 are a bimodal distribution where the two modes are separated by 1.5 nm [Fig. 7(b)]. For D1 to minimize overlay error back to this bimodal distribution, it is best for D1 to be positioned between the two modes of the C1/C2 distribution, i.e., split the difference between the systematic translation errors of the C1 and C2 distributions. If controlling systematic translation error within ± 0.5 nm is met (1.0-nm systematic translation error between prior patterns), then the single-layer capability could be relaxed to 3.0 nm to support a 3.5-nm overlay PA for D1 to the C1/C2 union. Thus there can be a trade-off between

the inherent single-layer overlay error capability and the systematic translation error control (between the multiexposed prior patterns) required to meet an overlay PA for an overlaying pattern to a composite prior pattern. For example, if inherent single-layer overlay error control is not good enough, a fab has the option of implementing translation overlay control limits between the prior patterns, e.g., using the example from Fig. 2(c) and Eq. (19) $(\overrightarrow{IP}_{C1} - \overrightarrow{IP}_{C2})$ must be below a certain control limit value, otherwise, the C2 exposure will be reworked so that new advanced process correction (APC) terms can be applied to the re-exposure of C2 in order to bring the $(\overrightarrow{IP}_{C1} - \overrightarrow{IP}_{C2})$ term within the control limits.

Setting design rules for multipatterned layers based on overlay PAs that are correctly determined using image placement and population-based statistics is critical. Without the proper statistical understanding, it can be concluded that the overlay capability cannot support a technology using multiple exposures, resulting in relaxed design rules and increased die areas. Of course, overly aggressive design rules will result in yield loss if the systematic translation error of prior patterns cannot be adequately controlled. As shown above, effective 3-sigma single-layer overlay capability needs to be tighter than the PA required by the design rules for an overlaying pattern minimizing overlay error to a prior pattern exposed with multiple exposures. We recommend setting the maximum systematic translation error between the exposures of a composite layer to be equal to 25% of the 3-sigma overlay PA of the overlaying layer to the composite prior pattern. As an example, if a 3-nm overlay specification is required for D1 to the composite layer formed with C1 and C2 in Fig 2(c), then the specification for the maximum systematic translation error between C1 and C2 needs to be <0.75 nm. If C2 to C1 systematic translation error were 0.8 nm then C2 would need to be reworked and APC correction applied to get the value within the specification.

4.4 Overlay Error of Overlaying Pattern to Union of “n” Prior Patterns

Sections 4.2 and 4.3 have illustrated the steps needed to derive fundamental equations that enable overlay error to

be calculated based on image placement errors and mean image placement error of the patterns involved when there are two prior patterns. This section expands some of the key equations to enable calculation of overlay error standard deviation when there are “ n ” prior patterns.

The first equation we expand is Eq. (14) for when there is no systematic translation error between the prior patterns. If rather than D1, C1, and C2, we use SEMI standard “2” for the overlaying pattern and then define the prior pattern PP_n , where n is the number of prior patterns, which overlaying pattern 2 needs to minimize overlay to, then the logic that enabled the derivation of Eq. (14) can be used to derive Eq. (21). Note that even though we have expanded to n prior patterns, the overlay error of overlaying pattern 2 to the n prior patterns remains the $\sqrt{2(\sigma_{ip})^2}$. This is because of the assumption that all n prior patterns have the no mean translations error between them and that they all have the same image placement error

$$\begin{aligned}\sigma_{2 \rightarrow PP_n} &= \sqrt{(\sigma_2)^2 + \frac{(\sigma_{PL_1})^2 + (\sigma_{PL_2})^2 + (\sigma_{PL_3})^2 + \dots + (\sigma_{PL_n})^2}{n}} \\ &= \sqrt{2(\sigma_{ip})^2}.\end{aligned}\quad (21)$$

Next, we expand the equations that calculate the standard deviation of the combined distribution when there is a mean translation error between the prior patterns. While Eq. (16) enables the calculation of the standard deviation of the union of two distributions, Eq. (22) is the more general equation for n distributions combining.²¹ Algebraically rearranging Eq. (22), the prior pattern image placement standard deviation, originally calculated with Eq. (17) for the case of a union of two prior patterns with equal counts, can be generalized to a union of n prior patterns [Eq. (23)]. Then Eq. (24) enables the overlay error standard deviation for any overlaying pattern “2” to any prior pattern union, regardless of whether there are systematic translation errors between the prior patterns, to be calculated [using the same set of assumptions that were used to derived Eq. (18)]

$$\begin{aligned}\sigma_{\text{union of } n \text{ distributions}} &= \sqrt{\frac{\sum_1^n N_n (\sigma_n)^2 + N_n (\mu_n - \mu_{\text{union of } n \text{ distributions}})^2}{\sum_1^n N_n}}\end{aligned}\quad (22)$$

$$\begin{aligned}\sigma_{\text{union of } PP_n} &= \sqrt{\frac{\sum_1^n N_{PP_n} (\sigma_{PL_n})^2 + N_{PP_n} (\mu_{PL_n})^2}{\sum_1^n N_{PP_n}} - (\mu_{\text{union of } PP_n})^2},\end{aligned}\quad (23)$$

$$\begin{aligned}\sigma_{2 \rightarrow PP_n} &= \sqrt{(\sigma_2)^2 + \frac{\sum_1^n N_{PP_n} (\sigma_{PP_n})^2 + N_{PP_n} (\mu_{PP_n})^2}{\sum_1^n N_{PP_n}} - (\mu_{\text{union of } PP_n})^2}.\end{aligned}\quad (24)$$

5 Impact of Measuring Overlay Error Back to Multiple-Prior Patterns

This section reviews the impact of measuring back to multiple-prior patterns and demonstrates that overlay metrology can report an overlay error lower than the actual overlay error due to averaging of the errors of the prior patterns. It should be noted that individual measurements of the overlaying pattern to the each of the prior patterns can be taken to understand the true overlay. However, overlay measurements of the overlaying pattern to the individual prior patterns require more complex APC algorithms be used to minimize overlay error. In addition, doing individual measurements to multiple-prior patterns increases the overlay metrology time. Therefore, unless both APC systems have been properly reconfigured for measuring an overlaying pattern to multiple-prior patterns independently and there is no concern with increased overlay metrology time, the overlay metrologist should aggregate, as will be described in Sec. 5.1, the multiple-prior pattern targets to enable the APC system to give proper feedback and minimize overlay metrology time. However, as will also be described, the value measured must be properly interpreted due to under measurement of the overlay error using the aggregate prior pattern methodology.

5.1 Measuring Overlay Error of an Overlaying Pattern Back to a Double Exposed Prior Pattern

If an overlaying pattern D1 minimizes overlay error back to two prior patterns C1 and C2 a target “CZ” can be defined on the overlay metrology tool (Fig. 8). Blossom^{22,23} or other overlay metrology targets that measure back to multiple-prior patterns simultaneously can be used to measure overlay error to this virtual target CZ_n , where the subscript n designates how many prior patterns are being aggregated into the CZ_n virtual target. By invoking a common reference grid for the prior patterns C1 and C2, the image placement error of the target CZ_2 ($\overrightarrow{ip_{CZ_2}}$) can be calculated from $\overrightarrow{ip_{C1}}$ and $\overrightarrow{ip_{C2}}$ [Eq. (25)]. The standard deviation of the image placement error of the CZ_2 target can be calculated using Eq. (26), if the image placement errors are random-independent variables

$$\overrightarrow{ip_{CZ_2}} = \frac{\overrightarrow{ip_{C1}} + \overrightarrow{ip_{C2}}}{2},\quad (25)$$

$$\sigma_{CZ_2} = \frac{\sqrt{(\sigma_{C1})^2 + (\sigma_{C2})^2}}{2}.\quad (26)$$

It should be noted the image placement error of CZ_2 is not directly measured. Rather Blossom (and other techniques) measure overlay error back to the composite multipatterned structure of C1 and C2 (CZ_2). Specifically, the overlay error of the overlaying pattern to the CZ_n target is measured. However, when developing the mathematics that explains the overlay error measured between an overlaying pattern and composite CZ_n target, starting with image placement is necessary.

The standard deviation of D1 to CZ_2 overlay error is derived by substituting D1 for pattern 2 and CZ_2 for pattern “1” in Eq. (9) [Eq. (27)]. Assuming $\sigma_{D1} = \sigma_{C1} = \sigma_{C2} = \sigma_{ip}$ (in other words, the standard deviation of image placement

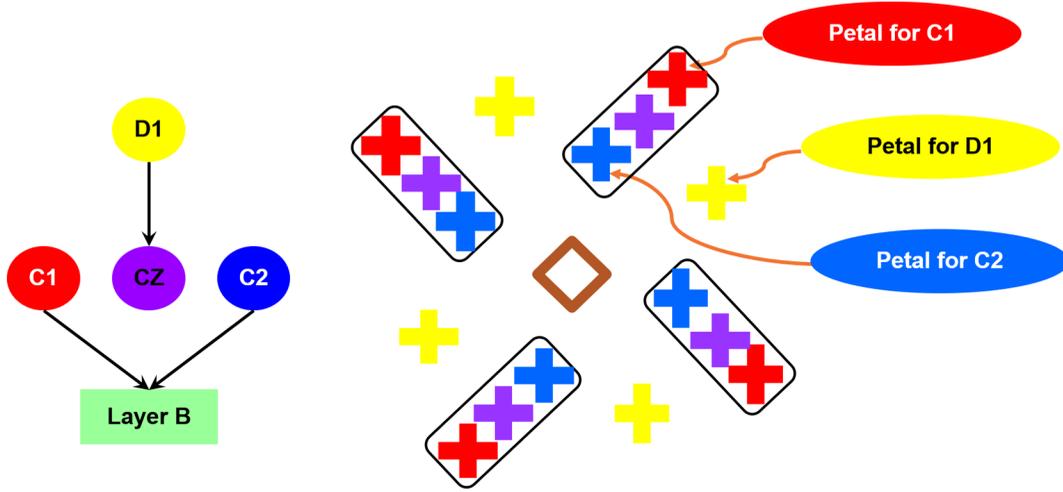


Fig. 8 Measuring overlay error back to prior patterns can be accomplished by measuring back to all prior patterns utilizing overlay metrology structures such as Blossom. The purple “CZ” target is a virtual target that is placed in this figure to assist in understanding why when measuring back to both C1 and C2 simultaneously with a multipattern target the true overlay error can be undermeasured.

error is the same for every pattern), the value of the standard deviation of D1 to CZ₂ overlay error is given by Eq. (28)

$$\begin{aligned}\sigma_{D1 \rightarrow CZ_2} &= \sqrt{(\sigma_{D1})^2 + (\sigma_{CZ_2})^2} \\ &= \sqrt{(\sigma_{D1})^2 + \frac{(\sigma_{C1})^2 + (\sigma_{C2})^2}{4}},\end{aligned}\quad (27)$$

$$\sigma_{D1 \rightarrow CZ_2} = \sqrt{(\sigma_{ip})^2 + \frac{(\sigma_{ip})^2 + (\sigma_{ip})^2}{4}} = \sqrt{1.5(\sigma_{ip})^2}. \quad (28)$$

5.2 Measuring Overlay Error of an Overlaying Pattern Back to a Triple Exposed Prior Pattern

Assume C1, C2, and C3 are the three exposures of a triple patterned layer. (Note: We do not show pattern C3 in Fig. 8 but it is evident that Blossom petals from the C3 or any other exposure could be added to the Blossom CZ_n target as appropriate.) If the overlaying pattern D1 minimizes overlay error back to the three prior patterns, C1, C2, and C3, and in metrology we define a new target, CZ₃, Eqs. (29) and (30) determine the image placement error and image placement standard deviation of CZ₃

$$\vec{ip}_{CZ_3} = \frac{\vec{ip}_{C1} + \vec{ip}_{C2} + \vec{ip}_{C3}}{3}, \quad (29)$$

$$\sigma_{CZ_3} = \frac{\sqrt{(\sigma_{C1})^2 + (\sigma_{C2})^2 + (\sigma_{C3})^2}}{3}. \quad (30)$$

Following the same logic as outlined for the two prior pattern case (Sec. 5.1), the sigma of D1 to CZ₃ overlay error is given by Eq. (31) and if $\sigma_{D1} = \sigma_{C1} = \sigma_{C2} = \sigma_{C3} = \sigma_{ip}$ the value of the standard deviation of D1 to CZ₃ overlay error us given by Eq. (32)

$$\begin{aligned}\sigma_{D1 \rightarrow CZ_3} &= \sqrt{(\sigma_{D1})^2 + (\sigma_{CZ_3})^2} \\ &= \sqrt{(\sigma_{D1})^2 + \frac{(\sigma_{C1})^2 + (\sigma_{C2})^2 + (\sigma_{C3})^2}{9}},\end{aligned}\quad (31)$$

$$\begin{aligned}\sigma_{D1 \rightarrow CZ_3} &= \sqrt{(\sigma_{ip})^2 + \frac{(\sigma_{ip})^2 + (\sigma_{ip})^2 + (\sigma_{ip})^2}{9}} \\ &= \sqrt{1.33(\sigma_{ip})^2}.\end{aligned}\quad (32)$$

5.3 Measuring Overlay Error of an Overlaying Pattern Back to a nth Exposed Prior Pattern

Equations (28) and (32) can be generalized to: if an overlaying pattern goes back to n prior patterns and all patterns involved have the same image placement error standard deviation (σ_{ip}), then the overlaying pattern to CZ_n overlay error standard deviation can be determined by

$$\sigma_{2 \rightarrow CZ_n} = \sqrt{(\sigma_{ip})^2 + \frac{(\sigma_{ip})^2}{n}}, \quad (33)$$

where n is the number of prior patterns.

5.4 Determining the Ratio of Measured to Actual Overlay Error for Multipatterned Systems

The ratio between metrology and actual overlay standard deviation can be exactly calculated for an overlaying pattern measuring back to a prior pattern patterned with multiple exposures. Specifically, using Eqs. (24) and (33), Eq. (34) can be derived to enable the determination of the ratio of measured to actual overlay error depending on the number of prior patterns. Equation (35) is the simplification of Eq. (34) when there are equal counts for each prior pattern and the image placement standard deviation for all prior patterns is the same value (σ_{ip}). Equation (36) is a further

simplification when the systematic translation errors for each prior pattern (μ_{PP_n}) are equivalent

$$\frac{\sigma_{2 \rightarrow CZ_n}}{\sigma_{2 \rightarrow PP_n}} = \frac{\text{measured overlay}}{\text{actual overlay}} = \frac{\sqrt{(\sigma_{ip})^2 + \frac{(\sigma_{ip})^2}{n}}}{\sqrt{(\sigma_2)^2 + \frac{\sum_1^n N_{PP_n} (\sigma_{PP_n})^2 + N_{PP_n} (\mu_{PP_n})^2}{\sum_1^n N_{PP_n}} - (\mu_{\text{union of } PP_n})^2}}, \quad (34)$$

$$\frac{\sigma_{2 \rightarrow CZ_n}}{\sigma_{2 \rightarrow PP_n}} = \frac{\text{measured overlay}}{\text{actual overlay}} = \frac{\sqrt{(\sigma_{ip})^2 + \frac{(\sigma_{ip})^2}{n}}}{\sqrt{(\sigma_{ip})^2 + (\sigma_{ip})^2 + \frac{\sum_1^n N_{PP_n} (\mu_{PP_n})^2}{\sum_1^n N_{PP_n}} - (\mu_{\text{union of } PP_n})^2}}, \quad (35)$$

$$\frac{\sigma_{2 \rightarrow CZ_n}}{\sigma_{2 \rightarrow PP_n}} = \frac{\text{measured overlay}}{\text{actual overlay}} = \frac{\sqrt{(\sigma_{ip})^2 + \frac{(\sigma_{ip})^2}{n}}}{\sqrt{(\sigma_{ip})^2 + (\sigma_{ip})^2}}. \quad (36)$$

Figure 9 illustrates the under measurement that grouped overlay metrology of multiple-exposed prior patterns has compared to the true overlay error as a function of the number of prior patterns using Eq. (36) (and the assumptions of equal image placement error of all prior patterns and no systematic translation error between the prior patterns). This is due to the point-by-point averaging of the image placement error of patterns that have been split into multiple exposures. As more prior patterns are used, the ratio of the measured overlay error to actual overlay error decreases. In other words:

- The overlay metrology results are giving the impression that overlay error is less than it really is.

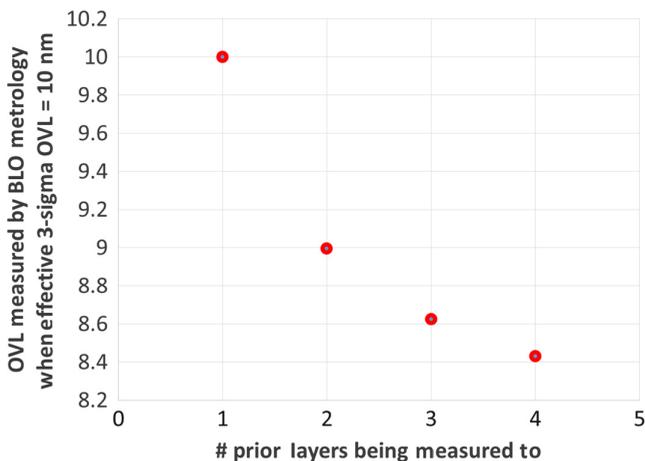


Fig. 9 The effects of Eq. (36) demonstrating that measured overlay error appears smaller as more prior patterns are used. The real overlay error stays the same at 10 nm.

- The difference between the actual and measured overlay grows as the number of prior patterns measured back to increases.

This is true no matter what the value is of the systematic translation error between the prior patterns. Specifically, whether there is no systematic translation error between the prior patterns or they have a systematic translation error of 10 nm, the point-by-point averaging will cause the overlay error measured to be less than actual overlay error. Indeed, when there is a systematic translation error between the prior patterns the measured overlay error will be even less representative of the actual overlay error than shown in Fig. 9. Of course, in such a case, Eq. (34) can be used to calculate the exact ratio.

Because of this under measurement of overlay error when measuring an overlaying pattern back to a multipatterned prior pattern, setting overlay error specifications that will determine whether a wafer will be reworked is more complex than that of single-layer cases. As described in Sec. 4.3 (see Fig. 7), the real overlay error between the overlaying pattern and the prior patterns increases as the systematic translation error between the prior patterns increases. Thus if using a grouped overlay metrology, both specifications of the overlaying pattern to the multiple-exposed prior pattern and the systematic translation error between the multiple exposures of a prior pattern need to be set so that the PAs can be supported.

6 Summary and Future Work

Methods discussed in the literature that look at “second order” overlay calculations and RSS the measured overlay errors to estimate indirectly controlled overlay error are not capable of estimating overlay error between an overlaying pattern and a prior pattern patterned with multiple exposures. Indeed, the prior literature did not address the problem of overlaying patterns that need to minimize overlay error back to a prior pattern composed with multiple exposures. Further, even though it is widely used, it was shown in Sec. 3 that the widely used RSS methodology for estimating overlay error that is indirectly controlled often over-estimates the overlay error due to not taking into the account that often the multiple exposures are correlated. New methods have been developed to estimate overlay error for multiple-exposed patterns (Sec. 4). These methods take advantage of going back to fundamental image placement error and population-based statistics. These methods allow for the proper estimation of overlay error between an overlaying pattern and a prior pattern that was exposed with one or more exposures. Specifically, a mathematical framework has been developed that can determine the impact of overlay error between exposures composing a multiple-exposed prior pattern, on the overlay error of a subsequent overlaying pattern, i.e., the impact of intralayer overlay error of the prior pattern on interlayer overlay error of the overlaying pattern that follows.

It was also shown that base single-layer process capability needs to be tighter than the PA of an overlaying pattern minimizing back to multiple-prior patterns with the specific amount of tightening directly related to the systematic translation error between multiple-exposed prior patterns. Because of this, systematic translation error specifications must be set appropriately between prior patterns to match

PAs (Sec. 4.3). Thus APC becomes an even more critical part of meeting overlay PAs. However, process variation coming from all sectors (not just lithography) must be minimized to enable APC to drive to the systematic translation error control needed.²⁴ Without this, semiconductor fabrication facilities will likely need to rework lots for systematic translation error even when absolute(mean) + 3*sigma overlay error is small.

Overlay metrology often undermeasures the overlay error for the case of an overlaying pattern measuring overlay error back to a prior pattern that was patterned with multiple exposures. The under measurement, compared to the actual overlay error, results from the aggregation of the prior patterns into a single-metrology target for the prior pattern. One of the benefits the aggregate prior pattern target enables is standard single-layer overlay error APC algorithms to be used with these multiple exposed prior pattern cases. However, overlay specifications for the semiconductor fabrication facility need to be adjusted accordingly due to the under measurement of overlay error. Finally, we again encourage others in the community to explore how systematics other than translation error affect the overlay error of an overlaying pattern to a multiple-exposed prior pattern.

7 Appendix A

7.1 Types of Overlay Error, Space Error, and EPE

When this paper refers to overlay, it always refers to centerline to centerline overlay. We use the descriptive term “space error” when considering the effect of combined CD and overlay error on the space between features. In the past, space error has been termed edge-to-edge overlay error.^{25–27} No matter what it is called the combined effect of CD and overlay error is well known to the design community as a key variable that not only affects space between two features but also design constructs that require overlap and intersect area (IA) between two shapes.²⁸ Edge placement error^{29–32} and relative edge placement error³³ are also terms that describe the effect of CD and overlay error together.

7.2 Overlay Metrology and Overlay Error Minimization

To measure overlay error of an overlaying pattern to a prior pattern, specific marks must be measured on a wafer. These marks will have structures from both the prior pattern and overlaying pattern. Usually, the prior pattern overlay metrology mark is an etched structure on a wafer. Because these metrology structures are surrogates to the actual device, the overlay metrology marks must be designed to be as close to the device of interest in terms of pattern size, pitch, and design density.

Direct overlay error minimization usually is used for minimizing the overlay error for layer interactions that are the most critical to preventing yield loss (see Sec. 7.3). Direct overlay error minimization refers to use of a process that utilizes at least two steps:

1. measuring overlay error between an overlaying pattern and a prior pattern and
2. using the measured overlay error and APC to minimize the overlay error of subsequent lots.

The overlay error minimization scheme is still considered direct even if the alignment on the scanner between the overlaying mask and the substrate is indirect. Specifically, sometimes setting the scanner to align to a different prior pattern (than overlay error is being minimized to) results in better direct overlay error minimization. Aung et al.³⁴ reviewed different alignment schemes and why sometimes better direct overlay error control results from indirect alignment.

Direct overlay error minimization should not be confused with indirect overlay error minimization, which refers to the overlay error between two patterns that is dependent on the overlay error of other patterns with direct overlay minimization. Usually, indirect overlay error minimization is used for layer interactions that are less critical to preventing yield loss (see Secs. 2, 7.3, and 7.5).

7.3 Yield Loss, Process Assumptions, and Rework Rate

When used in this paper, yield loss, PA, rework rate, and lot are defined as follows:

- Yield loss refers to chips that do not function or do not achieve performance requirements (speed, reliability, power consumption, etc.). Thus yield loss can come from defects but also from physical structures in the chip that are too small, large, close together (dielectric breakdown or electrical short) and/or far apart (electrical open). Combining CD error, line edge roughness and overlay error together can enable the determination of space error or minimum overlap area, both of which can directly impact yield.³²
- While overlay error can cause yield loss, image placement error by itself does not. Two hypothetical examples illustrate this point:
 1. If both an overlaying pattern and a prior pattern have the same large but identical image placement signatures (the vector fields of image placement are identical) then there will be no overlay error measured between the two patterns.
 2. If an overlaying pattern has no image placement error and the prior pattern has the large image placement signature, then there will be a large overlay error measured between the two patterns (even though the overlaying pattern had no image placement error).
- PAs are the specifications that a semiconductor fabrication facility needs to control for acceptable yield to be achieved, i.e., minimal yield loss. They are often documented as a target and a maximum standard deviation (usually as a 3-sigma value) of the distribution. An overlay PA documents the overlay requirements between an overlaying pattern and a prior pattern.³⁵ An overlay PA will typically have two distinct parts: (i) a target (which is usually zero) and (ii) a 3-sigma variation maximum. Refer to Sec. 2.4 of Ref. 32 for a more detailed discussion of PAs.
- Rework rate is the percentage of wafers that do not meet the PA(s) and are reworked. Rework typically involves removing the lithography film stack (e.g., the organic planarizing layer, inorganic hard mask,

and resist), then recoating and re-exposing the lithography layer. A wafer that does not meet its PA targets can be reworked and sent back through the process (usually with adjusted tool APC parameters) in order to achieve the PA. However, the rework process is ideally avoided due to both increased cost and degraded cycle time. For this reason, semiconductor fabrication facilities have targets for both yield loss and rework rate.

- A lot is a group of wafers that process together in a semiconductor fabrication facility. In this paper, we are concerned with overlay error and how it varies across/within:
 - a chip,
 - an exposure field, i.e., chip-to-chip variation if there are more than one chip per field,
 - a wafer (field-to-field variation), and
 - a lot (wafer-to-wafer variation).
- Lot-to-lot variation is also important to control for a semiconductor fabrication facility to have high yield.

7.4 Overlay Process Assumptions

The goal of the lithographic sector is to minimize the difference from the target of the average net overlay error measured for each lot and minimize the variation of overlay error ACFWL. The average overlay error is referred to as the translation error for the lot or lots. In this paper, when we say minimize overlay error, we are referring to both minimizing the difference from target ACFWL (having the average overlay error be zero) and minimizing the 3-sigma variation of overlay error ACFWL.

With single-exposure layers, the PA for overlay error of an overlaying pattern to a prior pattern could be met if the PA was larger than the on-product single-layer overlay capability. The on-product overlay capability was determined based on the overlay capability of the exposure tool under ideal conditions, error sources from the processes for the layers involved (wafer warping, stress variation, overlay metrology

induced error, APC induced error, etc.), and the rework rate the semiconductor fabrication facility was able to accept.

7.5 Overlay Error Control in Manufacturing

Overlay error between an overlaying pattern and a prior pattern can be measured and controlled. This minimization of overlay error is done by measuring the overlay error between two patterns and then using the measured overlay error to determine APC correction values to feed back for the next lot or even the current lot if the lot is reworked. However, yield of a semiconductor process not only is impacted by the value of the space variation and/or intersect area between the layers with direct overlay minimization but also by the choice of which prior pattern(s) to measure and minimize overlay error to using APC. When there is a choice of which prior pattern an overlaying pattern should be minimized back to, either design rule evaluations need to be made or experimental data needs to be obtained to determine the best minimization scheme for maximizing yield. Said another way: simply choosing the last pattern exposed, as the prior pattern for minimizing overlay error is not necessarily optimal for maximizing yield. This is true whether the process is using single- or multiple-exposure patterning. Figures 10 and 11 show examples of overlay minimization, with and without multiple-exposure patterning, where simply minimizing overlay to the last pattern exposed may not lead to the highest yield.

Figure 10 shows a via last dual damascene process that has no multiple-exposure patterning involved. In the via last case, the via can have its overlay error minimized to either the metal above or the metal below (MB). Note that even though the via is located between the two metal layers it was the last pattern exposed. To be specific, MB was patterned first, followed by metal above pattern which is a trench in the dielectric before via patterning. After the via pattern is patterned through the metal above trench, metallization of both the via and metal above patterns occurs in this via last process. As shown in Figs. 10(a) and 10(b), overlay error of the via to the MB along the x axis will decrease the IA between the via and MB. If the design rules are constructed so that this observation is true across all design

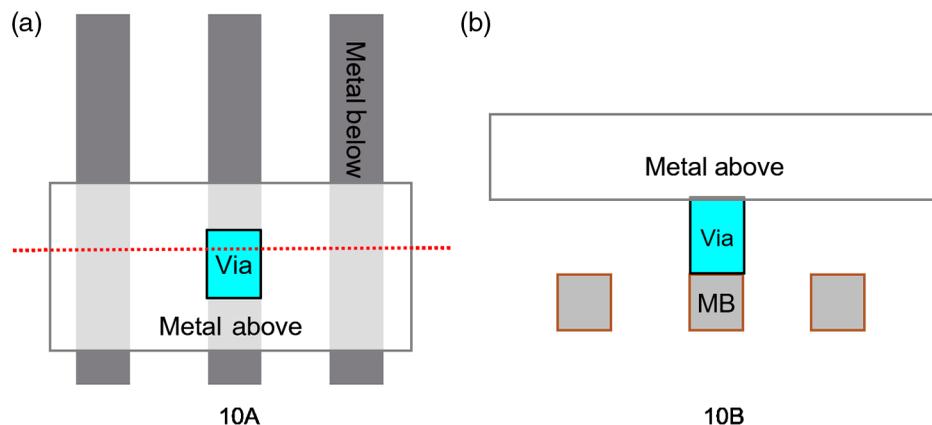


Fig. 10 (a) and (b) A via last dual damascene integration layout where the via should be aligned to the MB even though the metal above was the layer patterned immediately before the via. (a) The design layout and (b) a cross-sectional illustration where the cross-section location is noted by the dotted red line in (a). Note that even though the via was the last patterned layer it still connects (is between) the two metal layers.

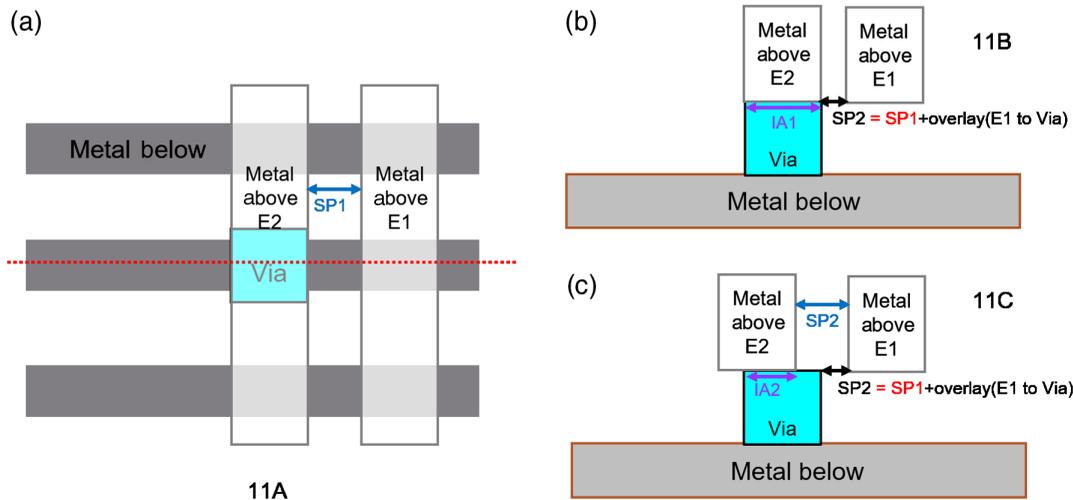


Fig. 11 (a) A via first dual damascene integration layout where the metal above is patterned with pitch split multiple-exposure patterning. (b) and (c) Cross-sectional illustration drawn at the dotted red line of (a). (b) If metal above E2 is minimized to the via layer the space between metal lines can be smaller than the target space. (c) Even though the metal above E1 exposure was the last patterned layer, if metal above E2 minimizes overlay error to the metal above E1 layer the IA between E2 and the via layer could be degraded.

constructs, then via overlay error should be minimized to the MB. Thus even though the metal above is the pattern exposed just before the via, in the via last process illustrated in Fig. 10, higher yield may be obtained by minimizing via overlay to the MB in the X orientation.

Figure 11 shows a via first scheme where multiple-exposure patterning is used for the metal above. In this case, the best choice for direct overlay minimization of the metal above second exposure (metal above E2) is less obvious. Specifically, overlay error of the second exposure (E2) can be minimized to the via layer or the first exposure of the metal above layer. To help illustrate the issues involved, cross-sectional illustrations are shown in Figs. 11(b) and 11(c). [Note: Fig. 11(a) has no overlay error between patterns while Figs. 11(b) and 11(c) illustrate two different possible overlay error situations as described below.] Metal above E2 minimizing overlay error to the via layer helps ensure that there is sufficient cross-sectional area between the via and metal above to carry needed current. However, Fig. 11(b) shows that if metal above E1 has an overlay error to the via that the space between the two metal lines (SP2) can become smaller than the target space (SP1). This can cause dielectric breakdown between the metal lines. Figure 11(c) illustrates the same via first dual damascene process where the first exposure of the metal above has the same overlay error relative to the via as in Fig. 11(b). However, in Fig. 11(c), the second exposure of metal above has its overlay error minimized to metal above E1. Minimizing overlay to the first exposure of the metal above helps maximize the amount of dielectric between the metal lines and thus minimize dielectric breakdown but can degrade via to metal above E2 IA as shown in Fig. 11(c), where IA2 is smaller than the target value [IA1 of Fig. 11(b)]. This smaller IA can lead to electrical opens. No matter which prior pattern is chosen for the metal-above-E2 overlay minimization, it is important to understand what the metal above E2 overlay error with the prior pattern that is not being directly minimized will be, so that design rules can be examined to

make sure that there are no failure modes. Depending on what is chosen, different statistical calculations need to be made to estimate the overlay error between other patterns that are not being directly minimized.³⁶ Sections 2 and 3 examine the different overlay error minimization possibilities and the statistical relationships for overlay error between what is minimized directly and what is minimized indirectly.

Acknowledgments

We thank Harry Levinson for his encouraging us to continue, write, and present this work. We thank Scott Halle, Andrew Brendler, and Chiew-Seng Koay for useful discussions. Finally, we thank SPIE and the Optical Microlithography Conference for the opportunity to present our work in this area.³⁷

References

1. L. Broussous et al., "TiN hard mask cleans with SC1 solutions for 64 nm pitch BEOL patterning," *Solid-State Phenom.* **219**, 209–212 (2015).
2. J. H. Chen et al., "56 nm pitch copper dual-damascene interconnects with triple pitch split metal and double pitch split via," in *IEEE Int. Interconnect Technol. Conf.*, San Jose, California, pp. 1–3 (2012).
3. L. W. Liebmann, K. C. Vaidyanathan, and L. Pileggi, "Multiple-exposure-patterning-enhanced digital logic design," Chapter 2, in *Design Technology Co-Optimization in the Era of Sub-Resolution IC Scaling*, p. 45, SPIE, Bellingham, Washington (2016).
4. C. Bencher et al., "22 nm half-pitch patterning by CVD spacer self alignment double patterning (SADP)," *Proc. SPIE* **6924**, 69244E (2008).
5. H.-Y. Tsai et al., "Sub-30 nm pitch line-space patterning of semiconductor and dielectric materials using directed self-assembly," *J. Vac. Sci. Technol. B* **30**, 06F205 (2012).
6. R. Gronheid and K. Ronse, "Lithography options for the 32 nm half pitch node and their implications on resist and material technology," *Proc. SPIE* **6827**, 68271V (2007).
7. T. Ebihara et al., "Beyond $k_1 = 0.25$ lithography: 70-nm L/S patterning using KrF scanners," *Proc. SPIE* **5256**, 985–994 (2003).
8. M. Hori et al., "Sub-40-nm half-pitch double patterning with resist freezing process," *Proc. SPIE* **6923**, 69230H (2008).
9. SEMI P18-92 (Reapproved 1104)—Specification for Overlay Capabilities of Wafer Steppers, <https://ams.semi.org/ebusiness/standards/SEMISTandardDetail.aspx?ProductID=211&DownloadID=5473>.

10. W. Arnold, M. Dusa, and J. Finders, "Manufacturing challenges in double patterning lithography," in *IEEE Int. Symp. Semicond. Manuf.*, Tokyo, pp. 283–286 (2006).
11. W. H. Arnold, "Toward 3 nm overlay and critical dimension uniformity: an integrated error budget for double patterning lithography," *Proc. SPIE* **6924**, 692404 (2008).
12. A. J. Hazelton et al., "Double patterning requirements for optical lithography and prospects for optical extension without double patterning," *Proc. SPIE* **6924**, 69240R (2008).
13. C. Lapeyre et al., "Impact of CD and overlay errors on double-patterning processes," *Proc. SPIE* **7274**, 72740W (2009).
14. J. D. Armitage and J. P. Kirk, "Analysis of overlay distortion patterns," *Proc. SPIE* **0921**, 207–223 (1988).
15. H. Ku, "Notes on the use of propagation of error formulas," *J. Res. Nat. Bur. Stand.* **70**(C), 263–273 (1966).
16. C. P. Ausschnitt and S. D. Halle, "Combinatorial overlay control for double patterning," *J. Micro/Nanolithogr. MEMS MOEMS* **8**(1), 011008 (2009).
17. H. J. Levinson and R. Rice, "Overlay tolerances for VLSI using wafer steppers," *Proc. SPIE* **0922**, 82–93 (1988).
18. C. J. Proglar, S. J. Bukofsky, and D. C. Wheeler, "Method to budget and optimize total device overlay," *Proc. SPIE* **3679**, 193–207 (1999).
19. E. Hendrickx et al., "Image placement error: closing the gap between overlay and imaging," *J. Micro/Nanolithogr. MEMS MOEMS* **8**(3), 033006 (2005).
20. J. P. T. Higgins and S. Green, Eds., *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*, Chapter 7, Table 7.7, The Cochrane Collaboration (2011).
21. S. C. Gupta and V. K. Kapoor, *Fundamentals of Mathematical Statistics (A Modern Approach)*, 10th ed., p. 3.11, Sultan Chand & Sons, Delhi (Revised Edition: August 2000, Reprint: 2002) (1975).
22. C. P. Ausschnitt et al., "Multilayer overlay metrology," *Proc. SPIE* **6152**, 615210 (2006).
23. N. M. Felix et al., "Smaller, smarter, faster, and more accurate: the new overlay metrology," *Proc. SPIE* **7638**, 76380Y (2010).
24. N. M. Felix et al., "Overlay improvement roadmap: strategies for scanner control and product disposition for 5-nm overlay," *Proc. SPIE* **7971**, 79711D (2011).
25. R. M. Booth, Jr. et al., "A statistical approach to quality control of non-normal lithographical overlay distributions," *IBM J. Res. Dev.* **36**(5), 835–844 (1992).
26. A. Starikov, "Metrology of image placement," Chapter 17, in *Handbook of Silicon Semiconductor Metrology*, A. C. Diebold, Ed., pp. 322–323, Marcel Dekker, Inc., New York (2001).
27. A. Starikov, "Special section guest editorial: control of integrated circuit patterning variance, part 2: image placement, device overlay, and critical dimension," *J. Micro/Nanolithogr. MEMS MOEMS* **15**(2), 021401 (2016).
28. C. Mead and L. Conway, *Introduction to VLSI Systems*, Addison-Wesley Longman Publishing Co., Inc., Boston, Massachusetts (1979).
29. J. K. Tyminski et al., "Lithographic imaging-driven pattern edge placement errors at the 10-nm node," *J. Micro/Nanolithogr. MEMS MOEMS* **15**(2), 021402 (2016).
30. N. Mohanty et al., "EPE improvement thru self-alignment via multi-color material integration," *Proc. SPIE* **10147**, 1014704 (2017).
31. J. Mulken et al., "Patterning control strategies for minimum edge placement error in logic devices," *Proc. SPIE* **10145**, 1014505 (2017).
32. A. H. Gabor et al., "Edge placement error fundamentals and impact of EUV: will traditional design-rule calculations work in the era of EUV?" *J. Micro/Nanolithogr. MEMS MOEMS* **17**(4), 041008 (2018).
33. J. Sturtevant et al., "Characterization and mitigation of relative edge placement errors (rEPE) in full-chip computational lithography," *Proc. SPIE* **9635**, 963505 (2015).
34. N. Aung et al., "Overlay control for 7 nm technology node and beyond," *Proc. SPIE* **10587**, 105870A (2018).
35. D. A. Chan et al., "Process for design of semiconductor circuits," US Patent 7, 861, 195 (2010).
36. A. J. Hazelton et al., "Achieving overlay budgets for double patterning," *Proc. SPIE* **7274**, 72740X (2009).
37. A. H. Gabor and N. M. Felix, "Overlay statistics for multiple exposure patterning," *Proc. SPIE* **10147**, 1014712 (2017).

Allen H. Gabor received his PhD in 1996 in materials science and engineering from Cornell University based on his work on block copolymer photoresists. He is a senior technical staff member at IBM. He has worked in the field of lithography at Arch Chemicals, GLOBALFOUNDRIES, and IBM. His work has included photoresist development, CD control, overlay minimization, and 193 dry, immersion, and EUV insertion. He is the author of more than 50 journal papers and holder of more than 30 patents. He currently serves on the program committee for the SPIE Extreme Ultraviolet (EUV) Lithography Conference and is a member of SPIE.

Nelson M. Felix received his BS degree from the University of Massachusetts, Amherst, in 2002, and his PhD from Cornell University in 2007, both in chemical engineering. Since joining IBM, he has managed various activities related to lithography infrastructure and control, and since moving to Albany, he has been focused on all elements to enable EUV lithography, including tooling, materials, and mask infrastructure. He is the current manager of the Foundational Patterning Group at IBM's Semiconductor Technology Research Center, Albany, New York. He has co-authored more than 90 papers and is the holder of 6 patents. He currently is a co-chair for the SPIE Extreme Ultraviolet (EUV) Lithography Conference and is a member of SPIE.