Optimal Feedforward Recipe Adjustment for CD Control in Semiconductor Patterning

Steve Ruegsegger, Aaron Wagner, Jim Freudenberg, Dennis Grimard

Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, Michigan 48109

Acceptable yields for nanofabrication will require significant improvement in CD control. One method to achieve better run-to-run CD control is through inter-process feedforward control. The potential benefits of feedforward control include reduced run-to-run post-etch CD variance, rework, and scrap. However, measurement noise poses a significant threat to the success of feedforward control. Since the stakes are high, an incorrect control action is unacceptable. To answer this concern, this paper will focus on how to properly use the available sensor measurement in a run-to-run feedforward recipe adjustment controller. We have developed a methodology based in probability theory that detunes the controller based on the confidence in the sensor's accuracy. Properly detuning the controller has the effect of filtering out the noise from the SEM. We will simulate this control strategy on industrial gate-etch data.

INTRODUCTION

The 1997 SIA Roadmap (7) suggests that 3σ CD control will need to be 10nm in 2003 (half of the current tolerance window) and 5nm by 2009. One method to achieve better run-to-run CD control is through inter-process feedforward control. Figure 1 shows a feedforward control system embedded into the patterning process. The lithography process has output X which is the input to the RIE process. The RIE has output Y. Disturbances D_{lith} and D_{rie} act on the outputs of the lithography and RIE, respectively. An in-line SEM is often employed in manufacturing systems for SPC on lithography CD. However, it may also be used for feedforward control. The measured photoresist (PR) CD is represented by M. The measurement also includes SEM disturbances, represented by D_{sem}. The feedforward controller adjusts the nominal RIE recipe in order to compensate for the estimated post-etch CD deviations $\hat{Y}$. The desired result of the RIE recipe adjustment is a reduction in the run-to-run variance of Y by rejecting D_{lith}. We call this strategy feedforward recipe adjustment (FFRA) control.

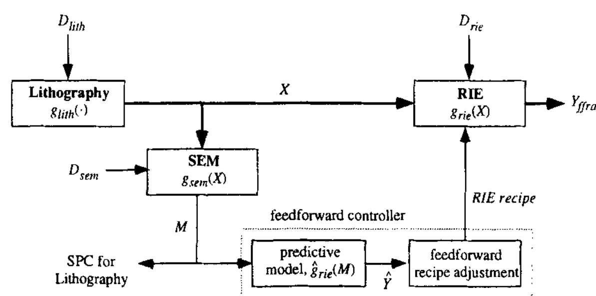

FIGURE 1. An inter-process feedforward controller implementation.

MOTIVATION

The potential benefits of feedforward control include reduced run-to-run post-etch CD variance, rework, and scrap. However, measurement noise D_{sem} poses a significant threat to the success of feedforward control. If the SEM noise is large enough, the measurement M will misrepresent the true PR CD X and the controller could command incorrect actions. Indeed, the variance of Y under feedforward control could actually increase! In most high-tech, high-cost manufacturing processes, incorrect control actions are unacceptable. The possibility of this scenario becoming reality is enough to prevent feedforward control from realization in manufacturing.

When a controller is subject to random measurement errors, compensation will only increase the variance of the process (4). Feedforward controllers can increase the output variance by making unnecessary or incorrect adjustments due to sensor noise. We will call this situation over-adjustment. The goal of this paper is to use variable estimation techniques to filter out the noise from the true underlying signal in order to avoid over-adjustment.

PREVIOUS WORK

Stoddard et al (10) implemented a feedforward and adaptive feedback controller in the manufacture of an on-chip capacitor. The goal was constant capacitance in the presence of varying dielectric thicknesses. Stefani et al (9) introduced a new eddy-current sensor and used it in a feedforward control strategy for controlling film thickness. They deliberately allowed a very wide input distribution in order to avoid over-adjustment. They state that the most important factor in their successful demonstration was having a very accurate and repeatable sensor measurement to feedforward. Leang et al (2) used a feedforward control strategy within a photolithographic stepper system. They recognize that feedforward control mechanisms are "... not well accepted in the semiconductor industry because of the high stakes involved.

CP449, *Characterization and Metrology for ULSI Technology: 1998 International Conference*
edited by D. G. Seiler, A. C. Diebold, W. M. Bullis, T. J. Shaffner, R. McDonald, and E. J. Walters
© 1998 The American Institute of Physics 1-56396-753-7/98/$15.00

A corrective action that worsens a process is not tolerated." Their approach is to only activate the feedforward controller when a SPC alarm triggered. Rietman (6) discusses a pre-production demonstration of a neural network that is used to regulate the resistance of vias between the first and second metal layers.

Only one of the above examples took actions to avoid over-adjustment in their feedforward control strategy. The others mentioned the possibility of over-adjustment, but commented that the variance of the sensor was known to be much less than the variance of the manufacturing process. This gave them confidence, at some level, that over-adjustment was not going to be a problem. However, this is not always the case.

SYSTEM VARIABLE DEFINITIONS

In order to investigate over-adjustment in feedforward control, the variables of the system (Figure 1) need to be defined. First, it will be assumed that all the random variables (RV) are deviations from target. We will also assume that the variables are independent, identically distributed (i.i.d.) zero-mean Gaussian. Finally, we assume that the noise terms are uncorrelated.

The lithography process is represented by the function $g_{lith}(\cdot)$, the SEM by $g_{sem}(X)$, and the RIE by $g_{rie}(X)$. These represent the nominal manufacturing processes without any noise. The deviation terms are defined to add noise to the outputs of the processes. The variance of D_{lith} is σ^2_{lith}, the variance of D_{sem} is σ^2_{sem}, and the variance of D_{rie} is σ^2_{rie}.

Figure 1 shows the lithography process is defined without an input deviation RV. Any disturbance contributed by the incoming wafers to the lithography can be represented by the disturbance D_{lith} without loss of generality. The output of the lithography system is simply:

$$X = D_{lith} \ . \tag{1}$$

Due to the simple setup of this system,

$$\mathrm{Var}[X] = \sigma^2_X = \sigma^2_{lith} \ . \tag{2}$$

The output RV of the RIE is dependent upon the type of control system implemented. Under nominal recipe conditions (no feedforward control), the output is represented by Y_{nom}. In order to define Y_{nom} and its variance, a useful model of $g_{rie}(X)$ needs to be specified. Guided by industrial data sets, we are going to express $g_{rie}(X)$ as a linear model:

$$\begin{aligned} Y_{nom} &= g_{rie}(X) + D_{rie} \\ &= aX + D_{rie} \ . \end{aligned} \tag{3}$$

The variance of Y_{nom} is:

$$\mathrm{Var}[Y_{nom}] = a^2\sigma^2_{lith} + \sigma^2_{rie} \ . \tag{4}$$

The PR mask CD is measured by a SEM. SEMs are designed to be very accurate measurement tools and are regularly calibrated to give linear, unity gain outputs with no offset over their range of operation. Therefore, the measured lot-mean PR mask CD deviation is modeled as:

$$\begin{aligned} M &= g_{sem}(X) + D_{sem} \\ &= X + D_{sem} \ . \end{aligned} \tag{5}$$

The variance of the SEM output is:

$$\begin{aligned} \mathrm{Var}[M] = \sigma^2_M &= \mathrm{Var}[X] + \sigma^2_{sem} \\ &= \sigma^2_{lith} + \sigma^2_{sem} \ . \end{aligned} \tag{6}$$

Equation 3 modeled the RIE as a linear system. Therefore, the predictive model of the RIE will use the same structure,

$$\begin{aligned} \hat{Y} &= \hat{g}_{rie}(\hat{X}) \\ &= \hat{a}\hat{X} \ . \end{aligned} \tag{7}$$

The predicted RIE output deviation, $\hat{Y}$, becomes the input to the FFRA controller. The controller outputs an adjustment to the nominal recipe in order to compensate for $\hat{Y}$. The result of this control action can be represented as subtracting the *predicted* nominal output deviation from the *true* nominal output deviation. Therefore, the RIE output under feedforward control is

$$Y_{ffra} = Y_{nom} - \hat{Y} \ . \tag{8}$$

There are three sources of error between the predicted output $\hat{Y}$ and the true value Y_{nom}:

1. the estimate of the true RIE input $\hat{X}$ does not represent the true input X exactly,
2. the model parameter estimate $\hat{a}$ does not represent the true model parameters a exactly, and
3. the linear model structure of $\hat{g}_{rie}(\cdot)$ does not capture the process $g_{rie}(\cdot)$ exactly.

Each error source should be minimized in order to increase the accuracy of the controller. There is significant work in all three areas in the statistics literature. Variable estimation can be found in most any statistics book, (1) for example. Techniques for modeling and parameter estimation can be found in (5). This paper will focus on removing error source #1 by applying variable estimation to FFRA control systems.

For the sake of illustration, the next section will define a simple estimation method that can result in over-adjustment. We will then derive a better estimator and compare results.

A NAIVE ESTIMATION

One particularly naive estimate of $\hat{X}$ is equating it to the reported SEM measurement,

$$\hat{X}_M = M \ . \tag{9}$$

If this were so, then the prediction of the RIE output, based on Equation 7, becomes:

$$\hat{Y}_M = \hat{g}_{rie}(\hat{X}_M) = \hat{a}M \ . \qquad (10)$$

Analogous to Equation 8, the FFRA output can be defined in terms of the nominal output and the control compensation $\hat{Y}_M$. Let Y_{naive} represent the FFRA output using the naive estimate,

$$\begin{aligned} Y_{naive} &= Y_{nom} - \hat{Y}_M \\ &= (aX + D_{rie}) - (\hat{a}M) \\ &= (a - \hat{a})X + D_{rie} - \hat{a}D_{sem} \ . \end{aligned} \qquad (11)$$

For the sake of simplicity, let us assume that the model parameter estimate is accurate ($\hat{a} = a$). Then,

$$Y_{naive} = D_{rie} - aD_{sem} \ . \qquad (12)$$

Since the noise terms are assumed independent, the post-etch variance becomes:

$$\text{Var}[Y_{naive}] = a^2\sigma_{sem}^2 + \sigma_{rie}^2 \ . \qquad (13)$$

Compare this FFRA variance to the nominal variance in Equation 4. One can see that this naive implementation of FFRA rejects the lithography disturbances *in exchange for* the measurement disturbances. This may or may not be a desirable thing to do.

Figure 3 plots the nominal variance and $\text{Var}[Y_{naive}]$ as a function of the measurement noise. Notice that if the measurement tool is perfect ($\sigma_{sem}^2 = 0$) then, under FFRA control, the input deviations would be compensated for exactly and only the RIE variance would remain. Also notice that under this naive implementation of feedforward control, $\text{Var}[Y_{naive}]$ can exceed the nominal variance. When the SEM variance is greater than the lithography variance, then feedforward control is *worse* than no control!

MMSE ESTIMATION

In contrast to the naive implementation described above, estimation theory can be used to define a better estimate of X. Classic signal processing techniques have a body of literature on estimating the value of an inaccessible RV in terms of the observation of an accessible RV. Since we are not ignorant about the RVs X or M, their expected behaviors can be used in the estimate of run-to-run x given m.

The problem statement is to find a $\hat{X}$ that minimizes the mean square error (MSE). The MSE is

$$\epsilon = E[(X - \hat{X})^2] \ . \qquad (14)$$

The $\hat{X}$ that minimizes the MSE is the minimum mean square error (MMSE) estimator. The MMSE estimator of X based on observing the RV M is the conditional mean (8):

$$X_{mmse} = E[X \mid M] \ . \qquad (15)$$

In general, solving the conditional expected value of the MMSE estimator is very difficult, except for the case of Gaussian RVs. Since we have assumed X and M are Gaussian deviations from target, the MMSE estimator can be calculated for our problem definition (3):

$$X_{mmse} = E[X] + \rho\frac{\sigma_X}{\sigma_M}(M - E[M]) \qquad (16)$$

where ρ is the correlation coefficient. This is the "optimal linear estimator" of X given M (8).

Consider the effect of ρ on the MMSE estimate. If ρ is zero (i.e. the RVs are uncorrelated), then the best estimate of X is its mean $E[X]$ and the measurement provides no useful information. When $\rho \neq 0$, the measurement M is included in the estimate with appropriate scaling. The correlation coefficient is calculated by:

$$\rho = \frac{\text{Cov}[X, M]}{\sigma_X\sigma_M} \ . \qquad (17)$$

The covariance of X and M is:

$$\begin{aligned} \text{Cov}[X, M] &= E[(X - \bar{X})(M - \bar{M})] \\ &= E[(X - \bar{X})(X + D_{sem} - \bar{X})] \\ &= E[(X - \bar{X})^2] \\ &= \text{Var}(X) = \sigma_X^2 \ . \end{aligned} \qquad (18)$$

Therefore, for our lithography and SEM setup,

$$\rho = \frac{\sigma_X}{\sigma_M} \ . \qquad (19)$$

The MMSE estimator becomes:

$$X_{mmse} = M\frac{\sigma_X^2}{\sigma_X^2 + \sigma_{sem}^2} \ . \qquad (20)$$

Let S represent the ratio of variances,

$$S = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_{sem}^2} = \frac{\sigma_X^2}{\sigma_M^2} \ . \qquad (21)$$

Note that S has the property

$$0 \le S \le 1 \ . \qquad (22)$$

Consider the affect of SEM noise on X_{mmse}. If there is no SEM noise ($\sigma_{sem}^2 = 0$), then $S = 1$ and $X_{mmse} = M$. That is to say, if the SEM is perfect, then the expected true input is, in fact, the measured value. When SEM noise exists, $S < 1$ and the expected true input X_{mmse} will be a fraction of the measured value M. Therefore, X_{mmse} will be closer to zero (target) than M. This is the de-tuning mechanism. As σ_{sem}^2 increases, $S \to 0$, which essentially turns off all control actions since $\hat{Y} = X_{mmse} \to 0$. The *a priori* information of lithography and SEM process variances is what detunes the controller and avoids over-adjustment.

Figure 2 shows a graphical example of this estimator. The lithography output has a standard deviation of 10 and the

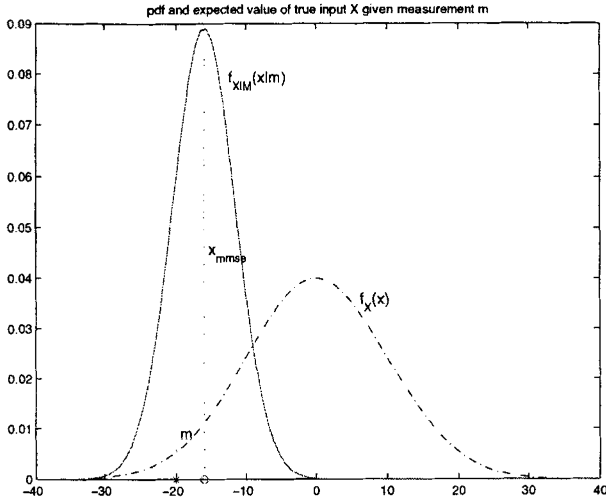

FIGURE 2. Distribution true input X given a measurement $M = m$.

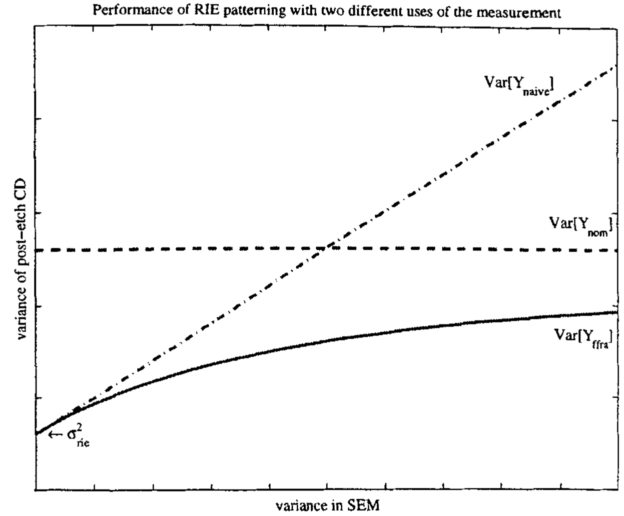

FIGURE 3. FFRA performance as a function of measurement noise.

SEM has a standard deviation of 5. A measured lot-mean CD deviation of $m = -20$ is given. The Gaussian distribution of the conditional RV $X|m$ is has expected value $E[X|m] = -16$ and variance of 20. Notice that x_{mmse} is closer to zero than the measurement m because of the disbelief in the measurement and *a priori* knowledge of the lithography distribution.

Using the MMSE estimator, the output of the RIE under FFRA becomes:

$$\begin{aligned} Y_{ffra} &= Y_{nom} - \hat{Y} \\ &= aX + D_{rie} - \hat{a}SM \ . \end{aligned} \tag{23}$$

The variance is calculated as:

$$\text{Var}[Y_{ffra}] = (a - \hat{a}\hat{S})^2\sigma_X^2 + \sigma_{rie}^2 + \hat{a}^2\hat{S}^2\sigma_{sem}^2 \ . \tag{24}$$

For simplicity, let us assume the model parameters are accurate ($\hat{a} = a$). A little algebra results in:

$$\begin{aligned} \text{Var}[Y_{ffra}] &= a^2 \frac{\sigma_X^2 \sigma_{sem}^2}{\sigma_X^2 + \sigma_{sem}^2} + \sigma_{rie}^2 \\ &= a^2 S\sigma_{sem}^2 + \sigma_{rie}^2 \ . \end{aligned} \tag{25}$$

Compare Equation 25 to Equation 4 and Equation 13. They all contain an independent σ_{rie}^2 term. Feedforward control does not reduce the inherent variance of the RIE process. The first term for all equations contain an a^2 representing the effect of the RIE process. For the nominal RIE process, the output variance includes lithography deviations. For the naive FFRA implementation, the lithography variance is exchanged with SEM variance. When the MMSE estimator is used in the FFRA implementation, the lithography variance is exchanged for a scaled SEM variance, $S\sigma_{sem}$. Since $S < 1$, only a fraction of the SEM variance enters into the variance of Y_{ffra}.

Consider the limits of the SEM variance to understand this FFRA variance. If the SEM variance is large, using the FFRA strategy will result in an output variance of:

$$\lim_{\sigma_{sem}^2 \to \infty} \text{Var}[Y_{ffra}] = a^2\sigma_X^2 + \sigma_{rie}^2 \ . \tag{26}$$

This is just the nominal output variance (Equation 4). The FFRA design will *turn off* the controller if the measurement noise is too big. On the other hand, if the measurement noise is small:

$$\lim_{\sigma_{sem}^2 \to 0} \text{Var}[Y_{ffra}] = \sigma_{rie}^2 \ . \tag{27}$$

The input deviations are being compensated for perfectly, and only the random noise of the RIE process remains. Obviously, feedforward control cannot reduce the output variance beyond the inherent RIE variance.

The variance using FFRA can be plotted against increasing measurement noise. Figure 3 shows that the variance during FFRA will not increase above the nominal variance. As the SEM variance increases, the FFRA design detunes the controller gains. Knowledge of the increased measurement noise decreases the measurement tool's credibility and the MMSE estimate reduces the amount of control authority.

RESULTS

We have simulated this FFRA methodology on a 0.35μm gate etch data set obtained from an industrial fab. The data set contains a pair of SEM measurements for each lot. The first measurement is the pre-etch PR mask CD. This corresponds to M. The second measurement is the post-etch gate CD. This corresponds to Y_{nom}. We will use these two data points to simulate the RIE output as if FFRA control had been used.

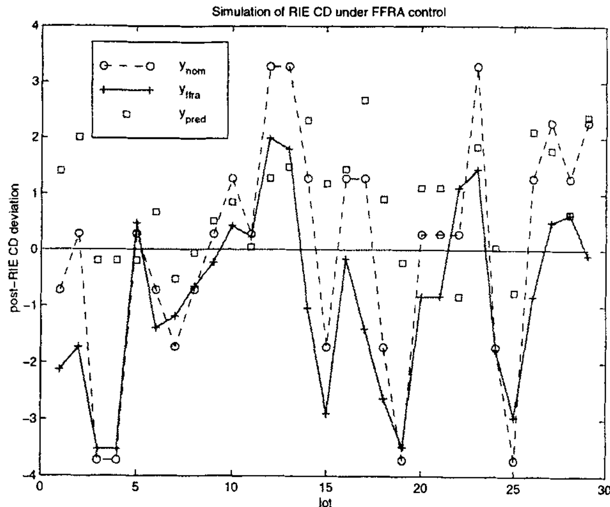

FIGURE 4. Simulated FFRA performances on normalized industrial data from a 0.35μm gate etch process.

First, we created the predictive linear model of M to Y_{nom} using linear regression (Equation 7). This calculated $\hat{a} = 0.73$. Next, we simulated the FFRA output by subtracting the predicted nominal output ($\hat{Y} = \hat{a}X_{mmse} = \hat{a}SM$) from the nominal output (Y_{nom}) as shown in Equation 23.

Figure 4 shows three data signals. The pre-etch measurements (M) and the RIE model were used to calculated the predicted nominal output ($Y_{pred} = \hat{a}X_{mmse}$) calculated from the pre-etch measurement and the RIE model. These are represented as '□'. The nominal etch deviations under no control (Y_{nom}) are shown by a dashed line and a 'o'. This is the second measurement from the data set directly. The simulated FFRA etch deviations (Y_{ffra}) are shown as a solid line and a '+'.

Notice that due to errors in the prediction, the controller does not always perform the proper action. For example, the measurement of the first lot predicted a CD above target, while the nominal CD was actually below target. Therefore, the FFRA simulated control action drove the output more negative. However, the MMSE estimation scaled the adjustment (by S) to avoid over-adjustment. Nonetheless, there are more corrections than improper adjustments. The standard deviation of the nominal output is $s[Y_{nom}] = 2.1$, while the standard deviation of the output with simulated FFRA control is $s[Y_{ffra}] = 1.6$. This is a reduction of 23%.

CONCLUSIONS

We have shown a methodology for proper integration of a sensor measurement into an inter-process feedforward controller. By using the MMSE estimator, over-adjustment is avoided and minimal variance is achieved.

Notice that this work is generic. It is applicable to many sensor and manufacturing processes. In fact, this work is currently being applied to lithography and RTP.

Future work will extend the use of MMSE estimation to another type of feedforward control strategy. Due to complexity and implementation issues, generating a unique recipe for each run may be undesirable. However, allowing the controller to select a recipe from within a pre-defined set of allowable, qualified recipes is sometimes acceptable. We call this control algorithm Feedforward Recipe Selection Control (FRSC). This approach will realize a portion of the FFRA benefits while minimizing the costs in complexity.

REFERENCES

1. E. Dougherty. *Probability and Statistics for the Engineering, Computing, and Physical Sciences.* Prentice Hall, 1990.
2. S. Leang, S. Ma, J. Thomson, B. Bombay, and C. Spanos. A control system for photolithographic sequences. *IEEE Transactions on Semiconductor Manfuacturing*, 9(2):191–207, May 1996.
3. A. Leon-Garcia. *Probability and Random Processes for Electrical Engineering.* Addison-Wesley Publishing Company, Inc., 2 edition, 1994.
4. J. MacGregor. A different view of the funnel experiment. *Journal of Quality Technology*, 22(4):255–259, Oct 1990.
5. J. Neter, W. Wasserman, and M. Kutner. *Applied Linear Statistical Models.* Richard D. Irwin, Inc., Homewood, IL, 2 edition, 1985. ch. 6.
6. E. Rietman. Multi step process yield control with large system models. In *Proceedings of the 1997 American Control Conference*, pages 1573–1574. American Control Conference, 1997.
7. The National Technology Roadmap for Semiconductors. Semiconductor Industry Association, 1997.
8. H. Stark and J. Woods. *Probability, Random Processes, and Estimation Theory for Engineers.* Prentice-Hall, Inc., 2 edition, 1994.
9. J. Stefani, K. Brankner, R. Jucha, W. Pu, and M. Graas. Open-loop predictive control of plasma etching of tungsten usin an in situ film thickness sensor. *SPIE Process Module Metrology, Control, and Clustering*, 1594:243–257, 1991.
10. K. Stoddard, P. Crouch, M. Kozicki, and K. Tsakalis. Application of feed-forward and adaptive feedback control to semiconductor device manufacturing. In *American Control Conference*, pages 892–896, June 1994.