

This document is downloaded from DR-NTU, Nanyang Technological University Library, Singapore.

Title	An efficient method for parametric yield gradient estimation(Published version)
Author(s)	Wu, Ting; Foo, Say Wei
Citation	Wu, T., & Foo, S. W. (1999). An efficient method for parametric yield gradient estimation. In Proceedings of the IEEE International Symposium on Circuits and Systems, (pp. 419-422). Singapore: National University of Singapore.
Date	1999
URL	http://hdl.handle.net/10220/5850
Rights	<p>© IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE. This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying this information are expected to adhere to the terms and constraints invoked by each author's copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder. This material is presented to ensure timely dissemination of scholarly and technical work. Copyright and all rights therein are retained by authors or by other copyright holders. All persons copying this information are expected to adhere to the terms and constraints invoked by each author's copyright. In most cases, these works may not be reposted without the explicit permission of the copyright holder.</p> <p>http://www.ieee.org/portal/site</p>

AN EFFICIENT METHOD FOR PARAMETRIC YIELD GRADIENT ESTIMATION

Ting Wu and Say Wei Foo

Department of Electrical Engineering
National University of Singapore, Singapore 119260
wut@vlsi.ee.nus.edu.sg
elefoosw@nus.edu.sg

ABSTRACT

A novel method to improve the yield gradient estimation in parametric yield optimization is proposed. By introducing some deterministic information into the conventional Monte Carlo method and fully utilizing the samples, it is possible to obtain yield gradient estimation with significantly smaller variance. The additional computation is almost negligible. Examples are presented to indicate the efficiency of this approach.

1. INTRODUCTION

Due to the inherent fluctuations in the manufacturing process, the performance of the mass-produced circuits varies from one to another. Thus not all the manufactured circuits necessarily meet the design specifications. Parametric yield optimization is a way to make the product more immune to these fluctuations by adjusting the nominal value of the designable parameters.

The yield can be described as the probability that a manufactured circuit has acceptable performance. Mathematically the yield for a nominal point $x^0 \in R^n$ is given by

$$Y(x^0) = \int_{R_d} \phi(x^0 + \theta) f_\theta(\theta) d\theta \quad (1)$$

where $f_\theta(\theta)$ is the joint probability density function (p.d.f.) of θ , the disturbance; R_d denotes the disturbance space and $\phi(\bullet)$ is an indicator function given by

$$\phi(p) = \begin{cases} 1 & p \in R_a \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

where R_a denotes the acceptability region in the parameter space R^n .

The yield optimization is then stated as

$$\max_{x^0 \in R^n} Y(x^0) \quad (3)$$

Ordinarily, gradient based approaches are efficient in solving the optimization problems. But in Monte Carlo method, only the estimation of yield gradient can be obtained and the variance of estimation is inevitable. It has been shown in applications that the efficiency of optimization is affected by the accuracy of the estimation to a great extent. In this paper, a method which provides a yield gradient estimator with smaller variance, and hence more accurate, is developed.

2. VARIANCE REDUCTION FOR YIELD GRADIENT

Assume that $f_\theta(\theta)$ in (1) is differentiable, the gradient can be expressed as an integral on the acceptability region as follows

$$\begin{aligned} \nabla Y(x^0) &= \int_{R_d} \phi(x^0 + \theta) \nabla f_\theta(\theta) d\theta \\ &= \int_{R_d} \phi(x^0 + \theta) \frac{\nabla f_\theta(\theta)}{f_\theta(\theta)} f_\theta(\theta) d\theta \\ &= \int_{R_{a'}} \frac{\nabla f_\theta(\theta)}{f_\theta(\theta)} f_\theta(\theta) d\theta \end{aligned} \quad (4)$$

where $R_{a'}$ denotes the corresponding area of the acceptability region in the disturbance space.

Since $f_\theta(\theta)$ is a p.d.f., the integral can be seen as the expectation of $\frac{\nabla f_\theta(\theta)}{f_\theta(\theta)}$ on the acceptability region, i.e.

$$[\nabla Y(x^0)] = E \left\{ \frac{\nabla f_\theta(\theta)}{f_\theta(\theta)} \right\}_{R_{a'}} \quad (5)$$

Therefore, the same set of samples distributed with the p.d.f. $f_\theta(\theta)$ for yield estimation can be employed here to give an unbiased estimate to $\nabla Y(x^0)$,

$$[\nabla Y\hat{(x^0)}] = \frac{1}{N} \sum_{i=1}^M \frac{\nabla f_\theta(\theta_i)}{f_\theta(\theta_i)} \quad (6)$$

where N is the sample number, $\{\theta_1, \theta_2, \dots, \theta_M\}$ is supposed in the acceptability region and $\{\theta_{M+1}, \dots, \theta_N\}$ outside the acceptability region.

2.1. Deterministic Information of the Equi-pdf Bounded Sub-region

Suppose the samples are statistically independent. Then the variance of the estimator in (6) is

$$\begin{aligned} \text{var}[\nabla Y\hat{(x^0)}] &= \frac{1}{N^2} \text{var} \left[\sum_{i=1}^M \frac{\nabla f_\theta(\theta_i)}{f_\theta(\theta_i)} \right] \\ &= \frac{1}{N^2} \sum_{i=1}^M \text{var} \left[\frac{\nabla f_\theta(\theta_i)}{f_\theta(\theta_i)} \right] \end{aligned}$$

$$= \frac{M}{N^2} \text{var} \left[\frac{\nabla f_{\theta}(\theta)}{f_{\theta}(\theta)} \right]_{R_a'} \quad (7)$$

which means the variance of the estimation is the sum of the variance of each sample inside the acceptability region divided by the square of the total sample number. The variance of a single sample is given by

$$\text{var} \left[\frac{\nabla f_{\theta}(\theta)}{f_{\theta}(\theta)} \right]_{R_a'} = \int_{R_a'} \left[\frac{\nabla f_{\theta}(\theta)}{f_{\theta}(\theta)} \right]^2 f_{\theta}(\theta) \mathbf{d}\theta - [\nabla Y(x^0)]^2 \quad (8)$$

Furthermore, if the acceptability region R_a' is partitioned into two disjoint subregions W and V , it can be shown that

$$\begin{aligned} \text{var} \left[\frac{\nabla f_{\theta}(\theta)}{f_{\theta}(\theta)} \right]_{R_a'} &= \text{var} \left[\frac{\nabla f_{\theta}(\theta)}{f_{\theta}(\theta)} \right]_W + \text{var} \left[\frac{\nabla f_{\theta}(\theta)}{f_{\theta}(\theta)} \right]_V \\ &\quad - 2\nabla Y(x^0)_W \nabla Y(x^0)_V \end{aligned} \quad (9)$$

where

$$\begin{aligned} \nabla Y(x^0)_W &= \int_W \frac{\nabla f_{\theta}(\theta)}{f_{\theta}(\theta)} f_{\theta}(\theta) \mathbf{d}\theta, \\ \nabla Y(x^0)_V &= \int_V \frac{\nabla f_{\theta}(\theta)}{f_{\theta}(\theta)} f_{\theta}(\theta) \mathbf{d}\theta \end{aligned}$$

and it is clear that $\nabla Y(x^0) = \nabla Y(x^0)_W + \nabla Y(x^0)_V$.

Hence if the yield gradient of a particular subregion (say W) is known deterministically (which means that its estimation variance is zero) and introduced into the yield gradient estimation for the whole acceptability region, the variance of the estimation will be absolutely reduced. According to (7) and (9), this reduction is achieved in two ways: by decreasing the variance of the individual sample by $\text{var} \left[\frac{\nabla f_{\theta}(\theta)}{f_{\theta}(\theta)} \right]_W$ and by reducing M , the number of random samples used in the summation.

There do exist a class of subregions whose yield gradient can be obtained with zero variance. In [5], it is indicated that the yield gradient can be expressed as a surface integral on the acceptability region boundary, i.e. along any direction \vec{z} , the gradient

$$\nabla_{\vec{z}} Y(x^0) = - \int_{\partial R_a'} f_{\theta}(\theta) \frac{\vec{z}}{\|\vec{z}\|} \bullet \mathbf{d}\vec{S} \quad (10)$$

where $\partial R_a'$ is the boundary of the acceptability region. If $f_{\theta}(\theta)$ is constant on the boundary, the above integral can be rewritten as

$$\nabla_{\vec{z}} Y(x^0) = -f_{\theta}(\theta) \frac{\vec{z}}{\|\vec{z}\|} \bullet \int_{\partial R_a'} \mathbf{d}\vec{S} \quad (11)$$

It is easy to prove that on an arbitrary closed surface ∂R_e

$$\int_{\partial R_e} \mathbf{d}\vec{S} = 0 \quad (12)$$

Thus for a region with the boundary which is a closed equi-pdf surface of $f_{\theta}(\theta)$, the yield gradient is constant zero.

So, if W is an equi-pdf bounded subregion of R_a' the yield gradient of the whole acceptability region is equal to that of the

remaining area, subregion V . Hence estimation can only be performed on subregion V . Compared with the conventional estimation in (6), only part of the acceptability region contributes to the variance. To establish a better estimator, we can find the largest equi-pdf bounded subregion in the acceptability region so that M and $\text{var} \left[\frac{\nabla f_{\theta}(\theta)}{f_{\theta}(\theta)} \right]_{R_a'}$ is minimum.

2.2. the Information of the Samples Outside the Acceptability Region

Let us define an inverse indicator function

$$\bar{\phi}(p) = \begin{cases} 0 & p \in R_a \\ 1 & \text{otherwise} \end{cases} \quad (13)$$

Since

$$\int_{R_d} f_{\theta}(\theta) \mathbf{d}\theta = 1 \quad (14)$$

the yield can be expressed as

$$Y(x^0) = 1 - \int_{R_d} \bar{\phi}(x^0 + \theta) f_{\theta}(\theta) \mathbf{d}\theta \quad (15)$$

and the yield gradient

$$\begin{aligned} \nabla Y(x^0) &= - \int_{R_d} \bar{\phi}(x^0 + \theta) \nabla f_{\theta}(\theta) \mathbf{d}\theta \\ &= - \int_{R_d} \bar{\phi}(x^0 + \theta) \frac{\nabla f_{\theta}(\theta)}{f_{\theta}(\theta)} f_{\theta}(\theta) \mathbf{d}\theta \\ &= - \int_{\bar{R}_a'} \frac{\nabla f_{\theta}(\theta)}{f_{\theta}(\theta)} f_{\theta}(\theta) \mathbf{d}\theta \end{aligned} \quad (16)$$

where \bar{R}_a' is the complement of R_a' . Correspondingly, the yield gradient estimator becomes

$$[\nabla Y(x^0)]' = -\frac{1}{N} \sum_{i=M+1}^N \frac{\nabla f_{\theta}(\theta_i)}{f_{\theta}(\theta_i)} \quad (17)$$

This estimator is independent of the conventional estimator in (6) but has the same expectation. Apart from the conventional method, it is only the samples outside the acceptability region that are taken into consideration for yield and yield gradient estimation. Although the estimator in (17) does not necessarily have a smaller variance than the estimator in (6), combining the two independent estimators in a proper way can lead to a new one having better performance.

Suppose that the variance associated with the estimator in (6) is σ_1^2 and that associated with the estimator in (17) is σ_2^2 . The linear combination of the two estimators

$$[\nabla Y(x^0)]'' = a[\nabla Y(x^0)] + b[\nabla Y(x^0)]' \quad (18)$$

where $a+b=1$, has the same expectation $[\nabla Y(x^0)]$ as $[\nabla Y(x^0)]'$ and $[\nabla Y(x^0)]''$, and its variance is $a^2\sigma_1^2 + b^2\sigma_2^2$. By setting

$$a = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}$$

$$b = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \quad (19)$$

the variance of the new unbiased estimator can achieve the minimum value $\frac{\sigma_1^2 \sigma_2^2}{\sigma_1^2 + \sigma_2^2}$, which is less than both σ_1^2 and σ_2^2 .

The conclusion on equi-pdf bounded subregion is also applicable in the complementary area of acceptability region, i.e. if the inner boundary of a subregion of \bar{R}_a' is equi-pdf, the integral in (16) on this subregion will be zero.

3. ALGORITHM

The following algorithm comprises the above two strategies:

- Generate a sequence of sample points from the p.d.f. $f_\theta(\theta)$.
- Examine the performance of each point and determine whether it meets the given specifications.
- Find the largest equi-pdf boundary inside which all the points meet the specifications and the smallest equi-pdf boundary outside which no point meets the specifications.
- Between the two boundaries, calculate two yield gradient estimates and their variances using the pass points and failure points respectively. Combine the two estimators according to (18) and (19) to obtain a new estimate.
- Using the new estimate as expectation, calculate the variances of the first two estimates and combine the two estimators again in the way of minimizing variance. Repeat this step until the new estimate converges.

Practical applications show that the convergence can usually be reached within five iterations.

4. EXAMPLES

4.1. Bandpass filter

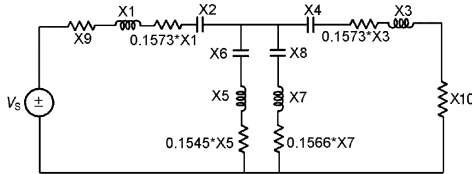
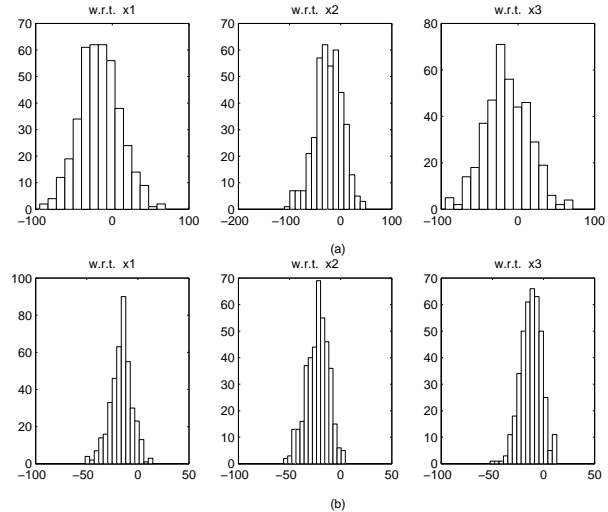


Figure 1: Schematic diagram of bandpass filter

The first example is a bandpass filter circuit^[9] shown in Fig. 1. The perturbances are assumed to be independent and follow Gaussian distribution as described in [9]. To illustrate the improvement achievable with the new method, the estimates of yield gradient at the initial nominal point given by [9] are computed using the new method and the conventional method respectively for the same set of samples. 400 experiments of yield gradient estimation are carried out with 100 samples in each experiment. Significant reduction of the estimation variances can be observed when the new method are used. Due to the limitation of space, only the distributions of the yield gradient with respect to the first 3 parameters are

shown in Fig. 2. The means and variances of the distributions are listed in Tab. 1.



(a) Conventional method

(b) New method

Figure 2: Distribution of the yield gradient estimation

4.2. Clock driver circuit

The second example serves to demonstrate the efficiency in yield optimization using the new yield gradient estimation method. The circuit for this example is shown in Fig. 3. It is a clock driver circuit discussed in [1] and [5]. The critical performance of interest is the clock skew defined as the delay between the rising clock signal of Node 6 and the falling inverted clock signal of Node 4 at 2V level. A tightened specification is set that the values of the clock skew must lie within $[-0.14, 0.14]$ ns. The same set of data in [1] except the initial nominal is employed.

Based on the new yield gradient estimation approach, the DFP (Davidon-Fletcher-Powell) method is used to solve the optimization problem with 100 samples for each iteration. The optimization improved the yield from the initial 19% to 92% after 3 iterations. Histograms of the skew values obtained from 100 simulation runs using the initial nominal design and optimized nominal design are shown in Fig. 4.

5. CONCLUSION

A new variance reduction method for yield gradient estimation is proposed. The method makes complete use of the available data of Monte Carlo samples. Two important concepts are presented. They are

- the yield gradient of the region with an equi-pdf boundary is deterministically zero;
- the samples outside the acceptability region can also provide an unbiased estimation for yield gradient.

Estimated gradient w.r.t.	Conventional method		Improved method	
	mean	σ	mean	σ
x1	-16.58	26.80	-16.48	10.61
x2	-23.74	26.91	-22.77	10.90
x3	-14.08	28.45	-12.24	10.28
x4	-17.27	25.54	-16.76	10.54
x5	-41.69	26.08	-38.15	12.01
x6	-23.72	27.43	-21.54	10.54
x7	-2.21	27.96	-1.40	10.43
x8	-4.75	25.41	-6.42	10.72

Table 1: Comparison of the distribution of the yield gradient estimation using conventional method and new method

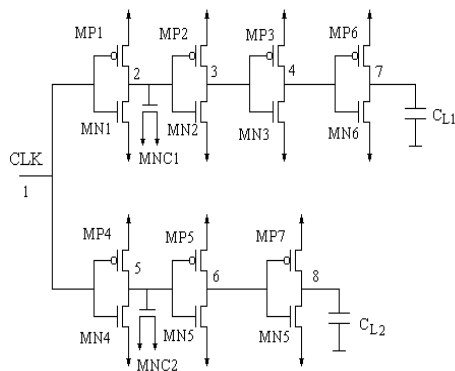
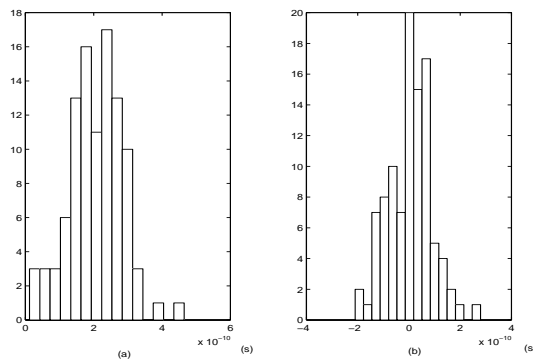


Figure 3: Clock driver circuit



(a) Before optimization
(b) After optimization

Figure 4: Distribution of clock skew

These concepts are applied to construct a more accurate estimator. Because the approach only attempts to make the effective use of the known information about the problem, no additional samples are required and the increase in computational load is minimal. The new gradient estimator with significantly reduced variance can lead to higher efficiency in the yield optimization.

References

- [1] J. C. Zhang and M. A. Styblinski, "Yield and Variability Optimization of Integrated Circuits", Kluwer Academic Publishers, 1995
- [2] R. Spence and R. S. Soin, "Tolerance Design of Electronic Circuits", Addison-Wesley Publishing Company, 1988
- [3] A. J. Strojwas, ed. "Selected Papers on Statistical Design of Integrated Circuits", IEEE Press, 1987
- [4] R. Y. Rubinstein, "Simulation and the Monte Carlo Method", John Wiley and Sons, 1981
- [5] P. Feldman and S. W. Director, "Integrated Circuit Quality Optimization Using Surface Integrals", IEEE Trans. on Computer-aided Design of Integrated Circuits and Systems, vol. 12, no. 12, pp. 1868-1879, Dec 1993
- [6] M. A. Styblinski and J. C. Zhang, "Orthogonal Array Approach to Gradient Based Yield Optimization", Proc. 1990 IEEE Int. Symp. Circuits and Systems, pp. 424-427, New Orleans, May 1990
- [7] M. A. Styblinski and A. Ruszczynski, "Stochastic Approximation Approach to Statistical Circuits Design", Electronics Letters, vol. 19, no. 8, pp. 300-302, 14th April 1983
- [8] D. E. Hocevar, M. R. Lighter, and T. N. Trick, "A Study of Variance Reduction Techniques for Estimating Circuit Yields", IEEE Trans. on Computer-Aided Design, vol. CAD-2, no. 3, pp. 180-192, July 1983
- [9] E. Wehrhahn and R. Spence, "The Performance of Some Design Centering Methods", Proc. 1984 IEEE Int. Symp. Circuits and Systems, pp. 1424-1438, Montreal, May 1984
- [10] S. W. Director, P. Feldmann and K. Krishna, "Optimization of parametric yield: A tutorial", Int. Journal of High-Speed electronics, vol. 3, no. 1, pp. 95-136, Mar. 1992