The Implementation of the Power Law Model in Silicon Oxidation Process Simulation

ZHI-MIN LING,* LUC H. DUPAS and KRISTIN M. DE MEYER†
IMEC, Kapeldreef 75, B-3030 Leuven, Heverlee, Belgium

The purpose of this work is to extend the power-law model to a semi-empirical oxidation model for process simulation. The proposed model benefits from a very simple model formulation, when compared to an existing model there are fewer parameter variables and simpler expressions of parameter dependence for a wide range of oxidation conditions while maintaining a very good overall fitting accuracy. The model covers dry oxidation for different pressures, temperatures, orientations, oxidation in a HCl/O2 ambient, doping enhanced oxidation phenomena, as well as wet oxidation for different pressure, orientation and temperature conditions. To develop the model, a strategy for parameter extraction, combined with optimal adaptation is proposed. This approach includes two steps. One is the determination of the parameter correlation matrix and the $\epsilon$-indifference range for each oxidation condition (e.g. temperature). The second step is the parameter adaptation within the $\epsilon$-indifference range to obtain a simple parameter variation over a wide range of oxidation conditions. As a consequence of this work, the proposed model is a possible replacement for the oxidation model of SUPREM-3 in process simulation. The optimized parameter extraction strategy can be applied to a large class of modelling problems.

Key words: Si, Si oxidation, power law model

1. INTRODUCTION

Over the past years, process simulation tools, e.g. SUPREM-3, ICECREM, PREDICT are more and more used in the design and optimization of both bipolar and MOS technologies. Since thermal Si oxidation is an essential process step in VLSI fabrication, all of these tools have the capability of Si oxidation simulation. Within these process simulation tools, the rate of SiO$_2$ growth on silicon has been described by the well-known linear-parabolic growth law (i.e. Deal-Grove model). The model proposer suggested that the oxygen species diffuse through the thermally formed oxide towards the SiO$_2$/Si interface where they react with Si atoms to form a new oxide layer. Therefore, this physical assumption results in a linear-parabolic law and can be written as:

$$X = a(t + t_o)^b$$

with $t_o = (X_o/a)^{1/b}$

$$X \propto (X^2 - X_i^2) + A(X - X_i) = Bt$$

(1)

Where $X$ is the oxide thickness, $t$ is the oxidation time, $B$ and $B/A$ are the parabolic rate and linear rate constant, and $X_i$ is the initial oxide thickness at time zero. This model is quite simple in formulation. However, as this linear-parabolic law is used to simulate the Si oxidation for different conditions, the model parameters $B/A$ and $B$ have to be modified according to these growth conditions. Therefore, the expression of the parameters as a function of different conditions (e.g. temperatures, pressures and orientations) of the linear-parabolic law can become quite complicated. In SUPREM-3, in order to maintain accuracy for dry oxidation in the thin oxide range (<200Å), extra exponential terms are added to the linear-parabolic law, and even look-up tables without any apparent physical meaning are used for oxidation in a HCl/O$_2$ ambient, due to the lack of knowledge in this case.

Reisman et al. proposed a power-law expression to model accurately both (111) and (100) silicon oxidation in dry oxygen for a wide temperature and pressure range. The model equation is quite simple and written as:

$$X = a(t + t_o)^b$$

with $t_o = (X_o/a)^{1/b}$

(2)

$X$ is the final oxide thickness for a given oxidation time $t$, $t_o$ is the time to grow an oxide of thickness $X_o$ which is already present on the silicon surface before the oxidation procedure starts. Nicollian et al. made an effort to place the power-law formulation on a solid theoretical footing. They argue that the thermal oxidation of silicon is surface reaction limited, and that the reaction rate is controlled by viscous flow of newly formed oxide to accommodate the volume expansion that occurs when silicon oxidises. The parameter $b$ is proposed to be a measure of the magnitude and rate of increase with time of the average oxide viscosity during the structural reconfiguration at oxidation temperature.

The difference between the physics underlying the linear-parabolic law of DG model and the power-law is obvious and illustrates that after more than 25 years of research the principles underlying the oxidation kinetics of silicon are still not very well understood. Therefore, we are not going to argue about the physical meaning of the parameter de-
pendences on growth conditions developed in this work, but, are mainly concerned with the development of a semi-empirical oxidation model efficiently usable for simulation.

Nevertheless, it is interesting to notice that the power-law can fit the oxidation data set from ultra thin to ultra thick oxide thickness ranges without using any extra terms. It is therefore obvious that the oxide thickness expression \( X = f(t) \) over the whole thickness range, is simpler for the power-law model than for the linear-parabolic law with an extra exponential term, since the latter needs a numerical integration to get the oxide thickness \( X \) for a given oxidation time \( t \). However, the extraction of the power-law parameters \( a \) and \( b \) in Eq. (1) is based on the criterion of 'best-fitting'. Very often, this 'best-fitting' based strategy results in an irregular variation of parameters over a wide range of process conditions. The irregular variation of parameter complicates the mathematical expression of the parameter dependence as a function of process conditions over a wide range. Thus the advantages (simplicity) of the power-law expression could be lost when a wide range of the process conditions is taken into account. Therefore, to have a simple semi-empirical oxidation model, suitable for process simulation tools, the power-law model should be extended to accommodate wider oxidation conditions, and the expressions of the parameter dependences on the different conditions (e.g., temperatures, pressures . . . ) should be as simple as possible, but still resulting in a good fitting accuracy. In order to reach these goals a strategy for parameter extraction combined with optimal adaptation is proposed.

In Sec. 2, the principles of the approach are described. The parameters extracted using the approach are formulated in Sec. 3. In addition, the most important experimental oxidation data sets published over last 20 years have been fitted by the power-law with the parameter dependences developed in this work. For comparison, the simulation results of new version of SUPREM-3 (version 8834, Dec. 1988) at these conditions are plotted together with the results of this work and are also presented in Sec. 3.

2. PARAMETER EXTRACTION PROCEDURE

The principles of the parameter extraction strategy used in this work can be found in a previous paper. The key points of this strategy are that, we consider the effect of propagation of experimental errors in the parameter extraction procedure, and that as such we do not necessarily adopt the 'best-fitting' parameter set for each individual data set (e.g., for a given oxidation temperature). Instead we consider if a particular acceptable parameter set (e.g., for a given oxidation temperature) can be combined together with other acceptable parameter sets at different conditions (e.g., oxidation temperatures) such that a simple formulation for the parameter variation over a whole range of process conditions (e.g., in the temperature domain) can be obtained. This means that the parameters for a specified condition are determined using a wider process condition range with concerns about the possibility of leading to a simple relation between the parameters as function of these conditions, rather than that they are extracted to achieve a 'best fitting' for each individual condition. The latter procedure i.e., 'best-fitting', not only very often results in an irregular variation of parameter over a wide range, but also the 'best fitting' parameter set values are not unique and depend on the actual fit criterion (absolute or relative error, weighting factors). This complicates the formulation of parameter variation as a function of growth conditions. It is possible, in our approach, that we might loose just a little bit of fitting accuracy for some individual data sets in a certain range, but the growth conditions dependences of the parameters, and thus the algorithms for the simulation become simpler. Moreover, if one considers the experimental error from run to run, one furnace to another, the fitting accuracy lost in our approach is not significant. The parameter extraction procedure includes two steps: one is the determination of the e-indifference range at each given condition (e.g., oxidation temperature), and the next one is the parameter optimal adaptation (adjustment) within each e-indifference range in a view of wide process conditions (e.g., temperature domain) to obtain a simpler parameter expression as a function of growth conditions. Obviously, this approach makes a compromise to some extend between the individual fitting accuracy and the global simplicity and transparency of the parameter dependence.

Determination of e-indifference range. The e-indifference range as defined in is a region in the \( n \)-dimensional parameter space for which the fitting error for each individual parameter set is smaller than \( \varepsilon \). How to obtain this range numerically for a given model and a given data set is explained in this section using a simplified example.

If there is only a native oxide present on the silicon surface before the oxidation procedure starts, \( i.e., X_0 = X_n \), the power-law expression in Eq. (2) can be approximated as:

\[
X = at^b + X_n = f(\hat{\theta}, t_n)
\]  

(3)

The oxidation procedure is assumed to start from a bare silicon surface with a native oxide, therefore, Eq. (3) can be used. Unfortunately, the value of native oxide thickness \( X_n \) differs from wafer to wafer depending on e.g. the precleaning procedure. To be simplified, in the following study, \( X_n \) is treated mathematically as a fitting parameter, rather as a real physical native oxide with a thickness determined by measurement.

The object function to be minimized in the parameter determination procedure is defined as:

\[
R(\hat{\theta}) = \sum_{\mu=1}^{N} \left( \frac{X_\mu - f(\hat{\theta}, t_\mu)}{X_\mu} \right)^2 / N
\]  

(4)

where, \( X_\mu \) and \( f(\hat{\theta}, t_\mu) \) are the measured and simulated (Eq. (3)) oxide thicknesses, respectively, \( \hat{\theta} \) is