

The Effect of Various Sputtering Parameters on Ta Phase Formation Using an I-Optimal Experimental Design

Charles S. Whitman
Lexmark International, Inc., Lexington, Kentucky

Abstract

Five process factors are varied to find conditions necessary for α and β phase formation in tantalum thin films deposited on SiC. These are: sputtering time, input power, pre-sputter etch time, preheat time at 250 C and sputtering temperature. An empirical model is developed which predicts the maximum or minimum amount of β phase possible over a large range of film thickness (~25 to ~2,000 nm). The maximum predicted (average) % β phase at the maximum sputtering temperature is only 8%, with 95% confidence bounds of [5%, 12%]. The other factors place a much lower restriction on β phase formation. Pure α phase is easily produced over a wide range of operating conditions. Only a weak relationship is found between film thickness and phase composition. The Ta film resistivity increases with the amount of β phase, in agreement with the literature.

Introduction

Tantalum thin films are used in a variety of applications. For example, Ta can be used as a diffusion barrier between Cu and Si [1], a gate electrode in MOSFET's [2], and as an x-ray absorber in x-ray masks [3] [4] [5] [6] [7] [8]. Further, Ta is often used as a protective overcoat above the resistor in inkjet devices. This last application is our principal concern.

Usually, Ta is sputter-deposited in a mixture of two phases: α and β. The β phase (body-centered tetragonal) of Ta is harder than the α phase (body-centered cubic) [9]. A harder coating over the inkjet resistor may improve its cavitation resistance [10]. The primary goal of this experiment was to find a combination of sputtering parameters to grow α or β phase Ta on SiC. Rather than varying only one factor at a time, five factors were chosen and varied simultaneously in order to find their individual effects *and* interactions between them. This was accomplished using a computer-generated I-Optimal experimental design [11].

The first step in generating an I-Optimal design is to choose a model to be fitted. In this case, a polynomial model was used to account for five main effects, ten two-factor interactions, five squared terms and one intercept for a total of 21 terms. The resulting regression equation has the form

$$Y = b_0 + b_1x_1 + b_2x_2 + \ldots + b_5x_5 + b_{12}x_1x_2 + b_{13}x_1x_3 + \ldots + b_{45}x_4x_5 + b_{11}x_1^2 + b_{22}x_2^2 + \ldots + b_{55}x_5^2 \quad (1)$$

where the x_i's are the levels of the process factors, and the b's are estimates of the effects of these factors that produce the estimated response, Y. Next, from all possible combinations of factors, the computer algorithm chose 24 trials which minimized the average variance of the predicted response. In this case, three factor and higher order interactions were assumed to be negligible.

Five factors were selected which have an influence on phase formation. The first two were sputter time and input power. Obviously, both will have a large influence on film thickness. Further, deposition rate and film thickness have been found to influence crystal structure [9] [12]. Thus, both of these factors were included.

Another factor was pre-sputter etch (PSE) time. Previous work has demonstrated that a PSE can enhance the growth of β-Ta [5]. A PSE may also promote adhesion between the Ta and SiC layers, improving the reliability of an inkjet heater.

The temperature of the substrate during film growth is known to affect the film's crystal structure [4]. A preheat step prior to Ta deposition could have a similar influence. It may also help "boil off" any surface contamination prior to deposition. Therefore, both substrate temperature and preheat time were included in this experiment. From a manufacturing viewpoint, it is easier to vary the duration of the preheat than the temperature. To maintain compatibility between Ta deposition and other production processes, a preheat temperature of 250 C was chosen. Due to manufacturability issues, neither bias voltage nor vacuum pressure were investigated.

Experimental

Each process factor can be considered one axis of a multidimensional cube. By varying the factors over a broad range, a large "experimental space" is created. A large space has two advantages. First, it improves the chances of finding some combination of factors which will produce the desired response. Second, it reduces the need for extrapolation outside the range after the experiment is completed. The ranges of the five process factors are proprietary. They have been coded between -1 (for "low"), and +1 (for "high").

All experimental trials were performed in the Varian 3290 in a random order. The different settings are shown in Table I. The first 24 trials constituted the original design. The added trials were used to confirm the empirical model. Runs #1, #13 and #24 had identical conditions. Responses from these runs can reveal trends from the beginning to the end of the experiment.

Table I. Experimental Design

Run Order	Time	Power	PSE	Preheat	Temp.
1	0.074	-0.074	0.031	-0.105	0.031
2	-1	-0.236	1	0.387	-1
3	1	0.756	1	-0.880	1
4	-1	0.244	-1	-1	1
5	1	-0.920	1	-1	-0.569
6	1	1	-1	-1	-0.203
7	1	-1	1	0.788	1
8	1	1	0.326	0.464	-1
9	0.019	-1	0.097	1	-1
10	1	-0.019	-1	1	0.097
11	-1	-1	-0.203	-1	-1
12	-1	1	0.178	-0.430	0.178
13	0.074	-0.074	0.031	-0.105	0.031
14	-1	1	-1	1	-1
15	-0.756	-1	1	-0.880	1
16	0.920	-1	-0.569	-1	1
17	-0.244	1	1	-1	-1
18	1	-1	-1	-0.093	-1
19	-1	0.056	0.112	1	1
20	-0.071	0.071	-1	-0.623	-1
21	-1	-1	-1	0.464	0.326
22	0.236	1	-1	0.387	1
23	-0.056	1	1	1	0.112
24	0.074	-0.074	0.031	-0.105	0.031
25	1	-0.471	0.5	-0.5	-1
26	-0.247	1	-0.75	0.25	0.75
27	1	-0.765	-1	-1	-1

The base pressure in the 3290 was 2×10^{-7} Torr. Sputtering was done after back filling with Ar to a pressure of 3 mTorr (the PSE was performed at 7 mTorr). All trials were performed separately, since only one recipe could be run at a time. Each (100) Si wafer had a thin film stack composed of a ~600 nm thermal oxide layer topped by a dual layer of ~440 nm silicon nitride (SiN_xH_y) and ~260 nm silicon carbide (SiC_xH_y). In each trial, at least three "dummy" wafers were run before the experimental wafers to promote removal of contamination from the Ta target. Each run consisted of two wafers, except run #21 which used six wafers, and runs #25 and #27 which used three wafers each.

The 3290 has four stations, all in the same chamber. Each station holds one wafer. After a treatment is completed, the wafer is transferred to the next station. In this experiment, the PSE was done in station 1 at -1300 V. Before the etch, the wafer was given a 10 sec. equilibration time at 0 V, followed by a 6.5 sec. ramp and a 3 sec. settle time. Stations 2 and 3 were used for the preheat. The preheat temperature of 250 C was accomplished using Ar heating. The 3290 has the capability of using an Ar "delay". If the preheat time was shorter than either the PSE or sputter time, then the preheat was delayed so that all stations finished simultaneously. The preheat time given in Table I was the time spent in each station, hence, the total amount of preheat time was twice that listed.

The wafer in station 4 received an additional 10 second preheat prior to sputtering to bring the wafer to its target temperature. However, if the desired sputtering temperature was lower than the preheat temperature, no attempt was made to let the wafer cool down to the temperature of interest. In production, this will increase throughput.

Three responses were observed: film thickness, sheet resistance (Rs), and % β-Ta. Film thickness was measured at the center of the wafer using an angle lap technique. Sheet resistance was determined with a four point probe.

Phase composition was measured using x-ray diffraction. Each wafer was scanned on a custom built vacuum path diffractometer using β-filtered chromium radiation from a sealed tube source operating at 35 kV and 40 mA. Slits used were DS = 1, SS = 1, and RS = 0.6 mm. The scan range (46 - 61, 2θ) was selected so as to pass through the <002> reflection of β-Ta (d=2.6635 A, 2θ = 50.94) and the <110> reflection of α-Ta (d=2.338 A, 2θ = 58.64). The fraction β phase was approximated by

$$\text{fraction } \beta \text{ phase} \sim I_\beta/(I_\alpha + I_\beta) \quad (2)$$

where I_α and I_β are the integrated intensity of the α and β phase x-ray peaks, respectively.

Results

Empirical Model

Using Strategy™ software, an empirical model was developed from the data which can be used to predict the responses described in the last section. Strategy™ provides estimates for the b's in equation (1) based on the observations from all the trials. A response, Y (film thickness, sheet resistance, or % β phase), can then be predicted for any desired levels of the process factors (sputter time, input power, etc.) within the ranges used in the experiment.

The region of experimental space which produces required film properties is called the "sweet spot". The regression equation provides a recipe for making films with the desired characteristics. The results of the regression reveal that the 95% confidence interval for some of the coefficients contain zero. Hence, the "actual" value could be zero. Even so, none of the coefficients were dropped from the model since they are the best estimates available.

Transformations were used to make the distribution of the residuals more nearly normal, and stabilize the variance. Thus, approximate confidence intervals can be assigned to predictions. A logarithmic transformation was used for the sheet resistance. The fraction β phase was transformed via the well known equation [13]:

$$transformed\ \%\ \beta\ phase = 2 \times \arcsin(p^{1/2}) \quad (3)$$

where p is the fraction β phase calculated with equation (2). A Box-Cox transformation [14] was used for the film thickness, which gave an exponent of -0.35. This result was then transformed again using a technique described by Kuehl [15]. The final result is given by

$$transformed\ film\ thickness = 159.843\ (1 - t^{-0.35})^{38} \quad (4)$$

where t is in angstroms.

Testing the Model

The model is computed using data from all over the experimental space. In this experiment, the space was cubical. Clearly, if all the data were from the surface of the cube, the model would have difficulty predicting behavior at the center. One test of the model was performed by removing the data of runs #1, #13 and #24 which were near the center. This new model predicted the average film thickness produced for these runs would be 509 nm with 95% confidence bounds of [309, 837] nm. Under the same conditions, the sheet resistance would be 0.46 [0.41, 0.51] Ω/sq., with an average % β phase of 25% [16%, 35%]. The observed averages for these runs were 709 nm, 0.44 Ω/sq., and 19% β phase, respectively. Since the observed averages were within the 95% confidence intervals, the statistical test was passed. Also, from a practical standpoint, the observed and predicted values are in relatively good agreement.

Runs #1, #13 and #24 provide a look at trends during the experiment. The average thicknesses of these runs were 734 nm, 722 nm and 671 nm, respectively. Similarly, the average sheet resistance was 0.49, 0.41, 0.41 Ω/sq. while the average amount of β phase in each run was 18.6%, 16.8%, and 22.1% (see Table II). The variation between the runs was small. This implies that extraneous factors had the same effect on the responses throughout the first 24 trials.

Model Predictions

The model was used to predict the factor settings needed to produce the desired responses. Using the GridSearch option on Strategy™, the experimental cube was "sliced" into 16ths and the model evaluated at the intersections. The settings of the factors which met the desired criteria were saved. Consider a 600 nm Ta film, a typical thickness used in thermal inkjet technology. Exploiting the data from the first 24 trials, GridSearch found the predicted maximum amount of α or β phase, for a predicted film thickness range of 540-660 nm. It is assumed that a 100% α film is equivalent to a 0% β film. Runs #25 and #26 were the initial sweet spots for maximum and minimum β phase. The model predictions were:

Run #25
predicted avg. thickness: 646 [505, 827] nm
predicted avg. sheet resistance: 1.33 [1.24, 1.42] Ω/sq.
predicted avg. % β: 95% [92%, 97%]
Run #26
predicted avg. thickness : 655 [524, 818] nm
predicted avg. sheet resistance: 0.31 [0.29, 0.33] Ω/sq.
predicted avg. % β: 0% [0%, 0.3%]

Table II. Average responses for each run.

Run	Avg. Thk. (nm)	Avg. Rs (Ω/sq.)	Avg. % β
1	734	0.49	18.6
2	47	17.1	30.0
3	1937	0.10	0
4	88	3.9	2.5
5	423	1.3	43.6
6	2144	0.12	48.7
7	409	0.51	0
8	1995	0.22	71.6
9	211	1.75	17.2
10	1446	0.18	22.1
11	46	61	14.5
12	106	2.6	0.4
13	722	0.41	16.8
14	120	3.5	14.6
15	70	4.1	0
16	367	0.54	0
17	846	0.76	74.3
18	401	2.8	72.9
19	69	4.2	0.7
20	664	1.5	83.8
21	26	20.1	0
22	1333	0.15	0
23	1089	0.20	0.7
24	671	0.41	22.1
25	911	0.54	48.7
26	684	0.31	0
27	625	2.2	93.1

From Table II, the observed average responses for run #26 agreed well with prediction. Thus, the minimum β–Ta sweet spot is acceptable, although the film thickness is a little high. The model predicts that a slightly lower sputter time would reduce the film thickness, yet preserve the pure α phase. According to the model, sputtering at the highest temperature allows the other factors to be varied over a wide range and still produce essentially pure α-Ta films. Thus, the α phase can be readily grown in production.

On the other hand, the maximum β phase (run #25) predictions were in poor agreement with observation. This is likely due to changes in extraneous factors in the manufacturing process. Because of this, three new sweet spots were chosen after including the new data. The one that yielded the largest amount of β phase was run #27. The new predictions were:

Run #27
predicted avg. thickness: 621 [569, 677] nm
predicted avg. sheet resistance: 1.89 [1.75, 2.05] Ω/sq.
predicted avg. % β: 86% [81%, 90%]

Run #27 had a fraction β phase of 93% [90%, 95%] and sheet resistance of 2.2 [2.1, 2.3] Ω/sq as noted in Table II. These results were higher than predicted. Since the goal is

maximum β phase, this result is welcome. Even so, it does not appear that a 100% β phase film can be grown under the film thickness restraint of 540 - 660 nm. In fact, the model predicts a nearly 100% β-Ta film would result in an average film thickness of 1615 [1530, 1700] nm. This film thickness is too large for inkjet applications for reasons of thermal efficiency.

The sputter time for run #27 was the maximum for the design and would be unacceptable in production. A more reasonable "coded" time of -0.316 further limits the amount of possible β phase. In order to maximize the fraction β-Ta under this time constraint, the temperature should be minimized. Further, maximizing the input power easily allows a film thickness of 600 nm. Figure 1 is a contour plot of predicted fraction β phase as a function of preheat and PSE, assuming the above values for the other factors. Figure 2 is the corresponding contour plot for the predicted film thickness.

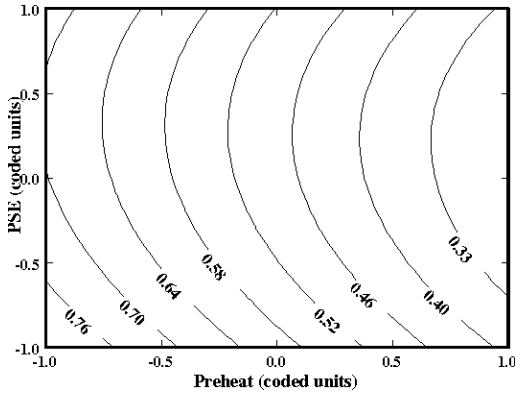

Figure 1. Contour plot of fraction β phase as a function of PSE and preheat time (sputter time = -0.316, input power = 1, and substrate temperature = -1).

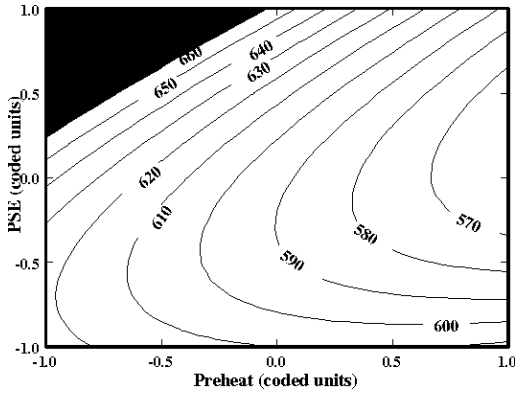

Figure 2. Contour plot of film thickness (nm) as a function of PSE and preheat time (sputter time = -0.316 s, input power = 1, and substrate temperature = -1).

From Figure 2, a film in the range of 540 - 660 nm is possible for most values of PSE and preheat, except for the top left corner, which is blackened. From Figure 1, the amount of β phase decreases with higher PSE times up to ~0.2. Above 0.2, the PSE increases the amount of β phase. This behavior is in agreement with Kondo [5]. Also from Figure 1, longer preheat times decrease the amount of β phase for any PSE time. This may be related to the removal of contamination from the surface [9]. The maximum fraction of β-Ta is 82% [77%, 87%], for minimum PSE and preheat.

Interpretation

Overwhelmingly, temperature has the largest influence on phase formation. If the temperature is set to the maximum, then the model predicts the maximum amount of β phase that can be grown is only 8% [5%, 12%], regardless of the settings of the other factors. When the temperature is minimized, the β phase is more easily grown.

The next most important factor is sputter time. Setting sputter time to -1, the maximum amount of β phase is predicted to be 63% [56%, 69%]. Low input power (-1) also limits the β phase composition to 81% [76%, 84%]. Higher sputter time and input power favor the formation of β–Ta. Nearly 100% β phase films can be produced over the full range of PSE and preheat times. However, sputter time, input power and substrate temperature must be adjusted to compensate, due to the presence of interactions.

While higher sputter time and input power can increase the amount of β phase, there is no direct relationship between film thickness and phase fraction. Figure 3 is a graph of fraction β phase vs. film thickness. Little trend is evident. Increased film thickness does not guarantee a larger fraction of β phase. Film thickness and phase are both *responses* to process conditions. A given film thickness or phase can be produced in a variety of ways. To produce the desired phase composition at a given film thickness, the best approach is to use the predictions of the empirical model. The predictions allow compromises to be made between the responses, as discussed in the last section.

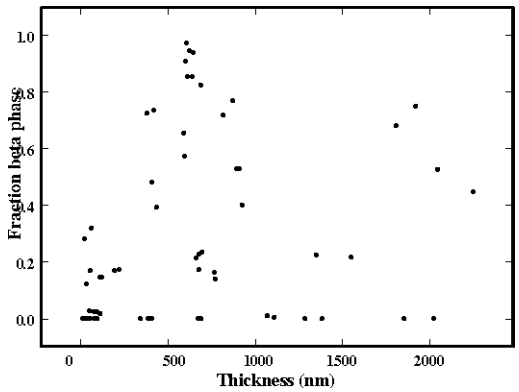

Figure 3. Scatter plot of fraction β phase as a function of film thickness.

The resistivity of the film should be a function of the fraction β phase present [16]. Resistivity is the product of sheet resistance and film thickness. Figure 4 reveals a quadratic relationship between these two responses. Three runs (#2, #11 and #21) had particularly high values of sheet resistance, even though the fraction β phase was low. Since these runs produced very thin films, it is likely the four point probe technique measured the relatively high

resistivity of the SiC layer below. Thus, the results from these runs are not shown.

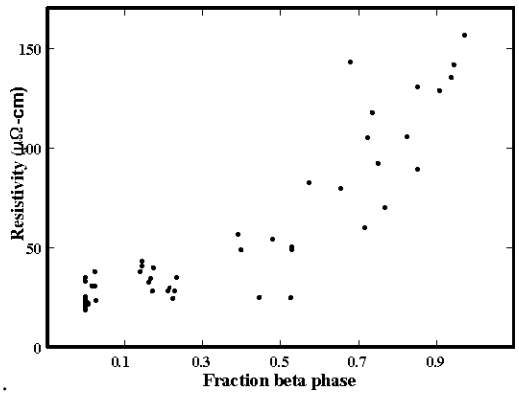

Figure 4. Scatter plot of film resistivity as a function of fraction β phase.

From Figure 4, a higher fraction β phase increases the resistivity, in agreement with the literature [16]. The resistivity varies from a low of 18.5 μΩ-cm for a 100% α-Ta film to a high of 156 μΩ-cm for a 97% β-Ta film. The empirical model can be used to predict the sheet resistance and fraction β-Ta for a given set of sputtering conditions. The phase fraction can be determined in the manufacturing line by measuring the sheet resistance and referring to Figure 4. The empirical model provides a map for process control.

Conclusions

Sputtering time, input power, temperature, pre-sputter etch time, and preheat time at 250 C were varied to find conditions necessary to produce α or β phase Ta films. An empirical model was developed which predicted the settings for the various factors to maximize or minimize the amount of β phase for a ~600 nm film. Using these settings, a film with 93% [90%, 95%] β phase was produced. Under conditions more suitable to manufacturing, the fraction β phase is predicted to be limited to 82% [77%, 87%]. While a ~100% β-Ta film is achievable, a film thickness of 1615 [1530, 1700] nm is necessary. In thermal inkjet, this film thickness is too high for both manufacturing and thermal efficiency reasons. On the other hand, a 100% α-Ta film was easily attained under the same film thickness constraint.

Prediction of phase composition is possible over a wide range of film thickness. Substrate temperature was found to be the strongest factor in controlling the crystal structure. Higher temperatures promoted growth of the α phase. To a smaller degree, low sputtering time and input power also limited the fraction of β phase. A direct relationship was found between fraction β phase and film resistivity. The larger the amount of β-Ta in the film, the higher the resistivity. Thus, sheet resistance can be used to monitor the phase composition in production. The empirical model serves as a map for maintaining process control.

References

1. Holloway, K. et al., "Tantalum as a diffusion barrier between copper and silicon", *Appl. Phys. Lett.*, **57**(17) Oct. 1990, pp. 1736-1738.
2. Shimada, H. et al., "Threshold voltage adjustment in SOI MOSFET's by employing tantalum for gate material", *Tech. Dig. IEDM*, 1995, pp. 881-884.
3. Yoshihara, T. et al., "Variation of internal stresses in sputtered Ta films", *JVST B* **11**(2), Mar/Apr 1993, pp. 301-303.
4. Yoshihara, T. et al., "Sputtering of fibrous-structured low-stress Ta film for x-ray masks", *JVST B* **12**(6), Mar./Apr. 1994, pp. 4001-4004.
5. Kondo, K. et al., "Stress stabilization of β-tantalum and its crystal structure", *JVST A* **11**(6), Nov/Dec 1993, pp. 3067-3071.
6. Kondo, K. et al., "Stress stabilization of tantalum absorbers on x-ray masks", *Microelectronic Eng.* **21**, 1993, pp. 75-78.
7. Jankowski, A. et al., "Internal stress minimization in the fabrication of transmissive multilayer x-ray optics", *JVST A* **7**(2), 1989, pp. 210-213.
8. Iimura, Y. et al., "Low-stress tantalum absorbers deposited by sputtering for x-ray masks", *JVST B* **7**(6), Nov/Dec, 1989, pp. 1680-1683.
9. Matson, D. et al., "High rate sputter deposition of wear resistant coatings," *JVST A*, **10**(4), Jul/Aug. 1992, pp. 1791-1796.
10. Horo, M.P., "Damage of thermal ink jet heaters by the collapsing vapor bubble", *SPIE* vol. **1079**, pp. 112-119, 1989.
11. Hardin, R. et al., "A New Approach to the Construction of Optimal Designs", *J. Stat. Planning and Inference*, **37**, 1993.
12. Hieber, K. et al., "Structural Changes of Evaporated Tantalum During Film Growth", *Thin Solid Films*, **90**, (1982) pp. 43-50.
13. Draper, N. and Smith, H., *Applied Regression Analysis*, John Wiley and Sons, 1981, p. 240.
14. Ibid, p. 225.
15. Kuehl, R., *Statistical Principles of Research Design and Analysis*, Duxberry Press, 1994, pp. 120-121.
16. Catania, P., et al., "Phase formation and microstructure changes in tantalum thin films induced by bias sputtering", *J. Appl. Phys.*, **74**(2), 15 Jul., 1993, pp. 1008-1014.

Biography

Charles S. Whitman graduated from Case Western Reserve University with a B.S. in Metallurgy and Materials Science in 1987. In 1992, he graduated from Northwestern University with a Ph.D. in Materials Science and Engineering. Since 1992, he has worked for Lexmark International, Inc. His job currently entails improving the reliability of the ink jet printhead, specifically the ink jet resistor. Additionally, he is pursuing a Masters degree in Statistics from the University of Kentucky.

e-mail address: whitman@lexmark.com