Integrated CMP Metrology and Modeling With Respect to Circuit Performance

Runzi Chang
Dissertation Talk
4/29/2004

Department of Electrical Engineering and Computer Science
University of California, Berkeley
Outline

- Motivation
  - Cu CMP process characterization and metrology
  - Model-based CMP process optimization
  - Impact of dishing on circuit performance
  - Conclusion
Interconnect delay is expected to strongly impact IC performance with aggressive scaling efforts

- Copper has been adopted for high performance
- CMP is enabling the Cu process

Source: The International Technology Road Map for Semiconductors: 2003
State-of-the-art Copper Interconnect

- **Back-End**
  - PVD, Etch – 1 Pass

- **BEOL**
  - Cu BEOL Interconnect
  - CVD, Etch, PVD, ECP, Cu CMP, Litho
  - 7-8 passes at 65nm

- **FEOL**
  - Contact
  - Transistor
  - EPI, Implant, RTP, Etch, CMP, Litho, CVD, Furnace – 1 Pass

Metal 9, M7, M8, M6, M5, M4, M3, M2, M1, Al Bondpad
Copper Damascene Process

- Cap/Low-k Deposition
- Damascene Etch
- Barrier/Seed

ILD → ECP → Cu CMP

Post-polish topography control (dishing, erosion) is key technical need
Previous Related Work

• Inter-layer dielectric (ILD) CMP modeling
  – Preston (1927) – the theory and design of plate glass polishing machine
  – Cook (1990) – chemical processes in glass polishing
  – Boning (1997) – analytic model for ILD thickness variation in CMP processes

• CMP metrology
  – Applied Materials: ISRM, Interferometry
  – KLA-Tencor: Scanning Electron Microscopy
Oxide CMP Metrology - Previous Project

Library-based scatterometry extracted profiles match AFM profiles and SEM pictures with 5nm precision.

Reference:
Previous Related Work (Cont’d)

• Copper CMP process modeling
  – Steigerwald (1997) – Chemical mechanical planarization of microelectronic materials
  – Gutmann (2002) – Book on Cu damascene process

• Interconnect variation analysis
  – Lin & Spanos (1998) – Circuit Sensitivity to interconnect variation
  – Nassif (2000) – Impact of interconnect variations on the clock skew of a gigahertz microprocessor
Challenges

• Lack analytical CMP models for copper damascene process
  – To estimate systematic metal and ILD loss
  – Copper dishing
    • Previous work used polynomial fitting to relate the metal line width to dishing
• Call for accurate, fast and non-destructive CMP metrology to develop, verify / refine the models
• Deep sub-micron scaling and wafer/die size increasing
  – Need tighter process control at 65nm technology node
  – Require precise electrical model of BEOL process variations on circuit performance
Outline

✓ Motivation

➢ Cu CMP process characterization and metrology
  • Model-based CMP process optimization
  • Impact of dishing on circuit performance
  • Conclusion
Test Pattern Design

- Designed E-test structures to measure and study dishing, where copper line width ranges from 0.4 to 5 microns
- Total Line length for each four point structure is about 4mm
  - Resistances are on the order of 100Ω depending on line width
Mask Design (I)

The difference in the resistance change percentage is a result of both erosion and dishing effects.

E-test cell

E-test cell

BCAM

4/29/2004
Mask Design (II)

- Mask has 7x7 array of E-test cells
- The effective pattern density for each cell is constant
- Resistance change is due to dishing metal loss
Process Flow

- Lithography and etching done at Berkeley microlab
- PVD and Cu CMP done at RPI Microlab
- About 30 wafers were polished with the planned DOE
Experimental Results

What are the causes of dishing?
Pad-Wafer Contact Mechanics

- Pad has a rough surface with random asperities
- The applied pressure is shared by pad asperities and the slurry film
- For normal platen speed, the hydrodynamic pressure of the fluid film is small, solid contact analysis is sufficient

Source: Motorola
Dishing Radius Concept

\[ dh = \frac{w^2}{4R_{dish}} \]

If \( w << R_{dish} \)

- \( R_{dish} \): effective radius of pad asperity;
- \( W \): metal wire width
- \( dh \): metal non-planarity due to dishing;

- Dishing “radius” is the ‘effective’ radius of pad asperities
- Function of asperity size, chemistry, particle size, pressure etc.

An array of cross-sectional SEM pictures qualitatively validated the dishing radius concept.
Concept Validation (Cont’d)

Surface profile of the post-CMP Copper structure (3µm metal line and 1 micron oxide spacing).
Wide lines R increase suffered most from dishing while narrow ones lost metal in corner rounding.
Profile Extraction using E-test Data

Assumptions:

a. Rounding is linewidth independent
b. Erosion is constant for a certain E-test structure (enforced by test pattern design)

\[ R = \frac{\rho \cdot L}{W \cdot t} \quad Y = \frac{1}{R} = \frac{W \cdot t}{\rho \cdot L} \]

Both \( \rho \) and \( L \) are const. Define \( y = Y \cdot \rho \cdot L \) we have

\[ y = y_0 - y_e - y_r - y_d \]

hence

\[ y_1 = W_1 t - W_1 dt - 2(1 - \frac{\pi}{4})r^2 - (aW_1) \cdot W_1 \]

\[ y_2 = W_2 t - W_2 dt - 2(1 - \frac{\pi}{4})r^2 - (aW_2) \cdot W_2 \]

\[ ... \]

\[ y_N = W_N t - W_N dt - 2(1 - \frac{\pi}{4})r^2 - (aW_N) \cdot W_N \]
Sample Extracted Results

Erosion=0.1µm, Corner rounding=70nm, Dishing Radius=40µm

\[
\frac{\rho \cdot \text{length}}{R} = \\
(w(t - dt) - 2(1 - \frac{\pi}{4})r^2 + \frac{wR_{\text{dish}}}{2} \sqrt{1 - \left(\frac{w}{2R_{\text{dish}}}\right)^2} - R_{\text{dish}}^2 \sin^{-1}\left(\frac{w}{2R_{\text{dish}}}\right)}
\]
Outline

✓ Motivation
✓ Cu CMP process characterization and metrology
  ➢ Model-based CMP process optimization
    • Impact of dishing on circuit performance
    • Conclusion
Identify Key CMP Performance Metrics

- Material Removal Rate (MRR)
- Selectivity
- Inter-layer dielectric erosion
- Metal dishing

And the key input parameters:

- Pressure
- Speed
- Slurry particle size
- Pad type
Linear Models

<table>
<thead>
<tr>
<th>Run No.</th>
<th>Pressure (3-6psi)</th>
<th>Speed 60-100rpm</th>
<th>Over-polish Time 15-30sec</th>
<th>W=5µm Dishing (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>1</td>
<td>30</td>
</tr>
<tr>
<td>2</td>
<td>-1</td>
<td>1</td>
<td>-1</td>
<td>22</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>-1</td>
<td>-1</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>20</td>
</tr>
</tbody>
</table>

\[
Dish = Dish_0 + \text{pressure} \times \text{Effect}_P + \text{speed} \times \text{Effect}_v + \text{time} \times \text{Effect}_t + \text{noise}
\]
\[
= 24.25 - 1.75 \times P - 3.25 \times V + 0.75 \times OT
\]

Similarly build linear models for each performance metric

\[
\text{Cost} = \sum_i w_i (\text{Target}_i - \text{Measured}_i)^2
\]

With known weights, process can be optimized by minimizing the cost function
Framework of CMP performance Optimization

1. Specification requirements
2. Identify CMP performance metrics
3. Identify key contributing Input parameters
4. Design of Experiments
5. Set weights for metrics
6. Identify Cost function
7. Model development
8. Process Optimization (optimizer)
9. Converging?

- Yes: Output
- No: Repeat steps 2-8
Taguchi’s Philosophy

• Use signal to noise ratio to measure and maximize performance

• Run a partial factorial set of experiments using orthogonal arrays
  – for every two columns all possible factor combinations occur equal times.

\[ S/N = -10 \log [MSD] \]

MSD refers to Mean Square Deviation of objective function
## Case Study Using Orthogonal Array

<table>
<thead>
<tr>
<th>Run No.</th>
<th>A (P(2-6psi))</th>
<th>B (S(80-100rpm))</th>
<th>C (slurry particle size 80-160nm)</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>w=5μm Dishing n(dB)</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>-30.9</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>-20.0</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>-18.1</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>-29.8</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>-26.0</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>3</td>
<td>1</td>
<td>-32.0</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>-21.6</td>
</tr>
<tr>
<td>8</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>-28.9</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>3</td>
<td>2</td>
<td>-24.6</td>
</tr>
</tbody>
</table>
Effects of Input Parameters

provides an intuitive, comprehensive view of the process
## Optimization Results

<table>
<thead>
<tr>
<th>Inputs</th>
<th>Starting Condition</th>
<th>Optimum Condition</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>setting</td>
<td>Dishing</td>
</tr>
<tr>
<td>A</td>
<td>A2</td>
<td>-29.3</td>
</tr>
<tr>
<td>B</td>
<td>B2</td>
<td>-25.0</td>
</tr>
<tr>
<td>C</td>
<td>C2</td>
<td>-24.8</td>
</tr>
<tr>
<td>Total Mean</td>
<td>-26.4</td>
<td>52.6</td>
</tr>
<tr>
<td>Real Output</td>
<td>20.8</td>
<td>426</td>
</tr>
</tbody>
</table>

The optimization process improves MRR by 10%, and reduces the erosion and dishing by >15%
Outline

✓ Motivation
✓ Cu CMP process characterization and metrology
✓ Model-based CMP process optimization
  ➢ Impact of dishing on circuit performance
• Conclusion
Model-Based Interconnect Performance Simulation

- Raphael™ simulations were carried out based on experimental results.
- RC delay was used as the metric to study the impact of dishing on interconnect performance.

Reference:
R. Chang, Y. Cao, and C. Spanos, “Modeling Metal Dishing for Interconnect Optimization,” IEDM 2003

2D Cross-sectional view and parameter definitions of global interconnect structure (line above one ground plane)
R, C vs. Linewidth with Dishing

\[ C_{total} = C_{ground} + 2 \times C_{coupling} \]
The optimal linewidth to achieve the minimum RC delay is around 4 µm.

Optimal range
RC Delay Sensitivity

- With dishing, RC delay becomes less sensitive to linewidth variation around the optimum width.
Efficiency of Process Improvement

Optimizing $W$ is critical in this region

$w = 4\mu m; t = 0.5\mu m$

$w = 4\mu m; t = 1\mu m$
Line-Splitting in Design

\[ w = \frac{w_{\text{total}}}{N}, \quad N \text{ is the number of lines}; \quad s = s_{\text{min}} = 0.5\mu m \]

- Line-splitting suppresses dishing, but sacrifices area and increases the fringing capacitance.
Gain in RC delay drops fast when splitting lines.

Optimal Splitting

\( N = 2 - 4 \)

\( W_{\text{total}} = 10 \mu m; \ R_{\text{dish}} = 40 \mu m \)
Optimization of Interconnect RC Delay

Optimal Splitting
$N=2-4$

- No Dishing
- $R_{\text{dish}}=40\mu m$
- $w_{\text{total}}=10\mu m$
- $t=0.5\mu m$
- $t=1.0\mu m$

RC Delay (ps/mm²)
N (Number of split lines)
Outline

- Motivation
- Cu CMP process characterization and metrology
- Model-based CMP process optimization
- Impact of dishing on circuit performance
- Conclusion
Conclusions

• Designed test patterns and implemented experiments to develop analytical dishing model, which well captured the systematic component of variations in Cu CMP
• Dishing radius concept was instrumental in process metrology
• Established multi-objective optimization framework in Cu CMP process
• Studied the interconnect performance and tradeoffs with dishing
  – Process improvement and/or layout modification is required when $R_{\text{dish}} < 50 \mu m$
  – Optimum line-splitting number is found to be 2-4 at design stage
Acknowledgements

• Professor Costas Spanos
• Prof. Cheung, Dornfeld – Thesis Committee
• Doyle, Neureuther, Poolla, Talbot - SFR/FLCC
• Dill, Gutmann - Mentors
• Berkeley, RPI, UCSB microlabs
• BCAM
• Family, collaborators and more