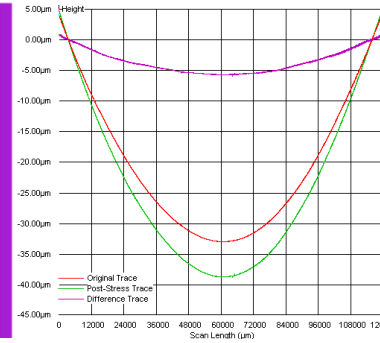


基于 KLA 探针式轮廓仪的薄膜应力测量

HRP® 系列和 Tencor® P 系列轮廓仪



介绍

随着器件尺寸不断缩小，表面翘曲度可能引发一些问题从而影响器件的正常功能。在半导体中，薄膜应力对半导体能带隙偏移、超导转变温度和磁各向异性等电子特性有直接影响¹。

在器件制造过程中，监控因薄膜沉积而产生的应力至关重要。在薄膜层面，应力通常会影响薄膜的附着力并产生晶格缺陷和表面再构，从而限制厚层薄膜的生长。在器件层面，应力的形成很少会直接导致良率下降，而是会缩短产品使用寿命。使用寿命缩短会导致每年在产品服务和保修方面花费数百万美元，这是一个重大问题。

结合成熟的应力计算理论模型与 KLA 探针式轮廓仪的精密测量，我们可以进行独立于材料与表面特性的精确薄膜应力测。

背景

应力无法直接测量；它是因薄膜沉积而产生的。薄膜沉积会导致基片弯曲并改变原始形状。基片的曲率半径可以通过测量基片翘曲度和挠度而获得。可以采用 G. Gerald Stoney 开发的用于测量薄膜应力的悬臂梁技术，通过比较薄膜沉积前后的曲率半径变化来估算应力²：

$$\sigma = \frac{1}{6R} \frac{E}{(1-\nu)} \frac{t_s^2}{t_f} = \frac{1}{6} \frac{E}{(1-\nu)} \frac{t_s^2}{t_f} \left[\frac{1}{R_f} - \frac{1}{R_s} \right]$$

其中

$E/(1-\nu)$ = 晶圆的弹性常数

t_s = 晶圆厚度

t_f = 薄膜厚度

R = 曲率半径

R_s = 衬底的曲率半径

R_f = 带有薄膜的衬底的曲率半径

E = 晶圆（衬底）的杨氏模量

ν = 泊松比

应力测量技术

HRP® 和 Tencor® P 系列轮廓仪可在无需拼接的情况下测量衬底的整片直径，可提供基片翘曲度和挠度的详细数据。首先收集薄膜沉积前和薄膜沉积后的扫描图像，然后通过计算由于薄膜沉积而引起的基片曲率半径变化，即可计算应力。

HRP 系列和 Tencor P 系列模型为用户提供两种应力计算算法： n 阶多项式和 13 点最小二乘拟合。前代模型仅提供一种方法：13 点最小二乘拟合。在多项式拟合程序中，用户可以指定使用 5 阶、6 阶或 7 阶多项式来计算曲率半径。 n 阶多项式有 $n+1$ 个系数：

$$y = C_0 + C_1x + C_2x^2 + \dots + C_nx^n$$

为了确定这些系数，考虑使用 3 阶多项式的示例：

$$y = C_0 + C_1x + C_2x^2 + C_3x^3$$

一共需要四个方程式来确定这些系数。可以简单地将以上方程式乘以 x^3 、 x^2 、 x 和 1（未知数的系数），生成这些方程式。生成的四个方程式如下：

$$x^3y = C_0x^3 + C_1x^4 + C_2x^5 + C_3x^6$$

$$x^2y = C_0x^2 + C_1x^3 + C_2x^4 + C_3x^5$$

$$xy = C_0x + C_1x^2 + C_2x^3 + C_3x^4$$

$$y = C_0 + C_1x + C_2x^2 + C_3x^3$$

使用 Crout 法求解联立方程式，即可确定这四个系数，并获得将高度表示为位置函数的多项式。曲率半径由多项式的二阶导数确定³。

13 点最小二乘拟合合法的步骤包括将局部数据拟合成弧线，并根据这些弧线的局部曲率计算平均曲率。由于这种方法多次使用相同的数据点，因此更易受到噪声和震动信号影响，可靠性不佳。

最小二乘拟合算法中剔除了所收集数据中前 5% 和后 5% 的数据。其余数据被分为三个长度为 0.3L 的数据段，其中 L 为扫描长度。通过计算第 1~13，2~14 直至 N-12~N 个点的局部曲率（其中 N 为数据段中数据点的总数），可确定局部的曲率半径。每个数据段的平均曲率是局部曲率的均值。

局部曲率的计算方法需要 2 阶多项式（其中二阶导数用于计算曲率）。多项式的一般方程式如下：

$$y = a_0 + a_1x + a_2x^2$$

该方程式的期望值为：

$$\hat{y} = a_0 + a_1x + a_2x^2$$

通过以下方程式将残差 $(y_i - \hat{y}_i)$ 平方和最小化：

$$\partial(\sum_{i=1}^{n+12} (y_i - (a_0 + a_1x_i + a_2x_i^2))^2) / \partial a_0$$

$$\partial(\sum_{n=1}^{n+12}(y_i - (a_0 + a_1x_i + a_2x_i^2))^2) / \partial a_1$$

$$\partial(\sum_{n=1}^{n+12}(y_i - (a_0 + a_1x_i + a_2x_i^2))^2) / \partial a_2,$$

其中 n 为第一个数据点。这些方程式进一步简化为：

$$13a_0 + a_1 \sum_{n=1}^{n+12} x_i + a_2 \sum_{n=1}^{n+12} x_i^2 = \sum_{n=1}^{n+12} y_i$$

$$a_1 \sum_{n=1}^{n+12} x_i^2 + a_2 \sum_{n=1}^{n+12} x_i^3 = \sum_{n=1}^{n+12} x_i y_i$$

$$a_0 \sum_{n=1}^{n+12} x_i^2 + a_1 \sum_{n=1}^{n+12} x_i^3 + a_2 \sum_{n=1}^{n+12} x_i^4 = \sum_{n=1}^{n+12} x_i^2 y_i$$

然后，我们可以使用矩阵行列式或简单替代法求解系数的这些方程式，然后计算曲率。

收集应力数据

对于精确应力测量，使用针对应力配置的通用载台或应力载台非常重要。应力载台具有三个用于支撑样品的支撑销钉以及两个或三个用于辅助定位晶圆的定位销（有缺口的晶圆使用两个定位销，有平边的晶圆使用三个定位销）。*需要注意的是，支撑销钉和定位销的位置需要根据样品大小进行调整。*支撑销钉将样品保持在适中位置，通过均匀分布样品重量，防止 (a) 载台形状和 (b) 重力对晶圆形状产生影响。定位销确保样品被放置于一个可复现的位置，从而提高准确性、重复性和安全性。图 1 显示了用于 8 英寸（200 毫米）晶圆的应力定位载台。

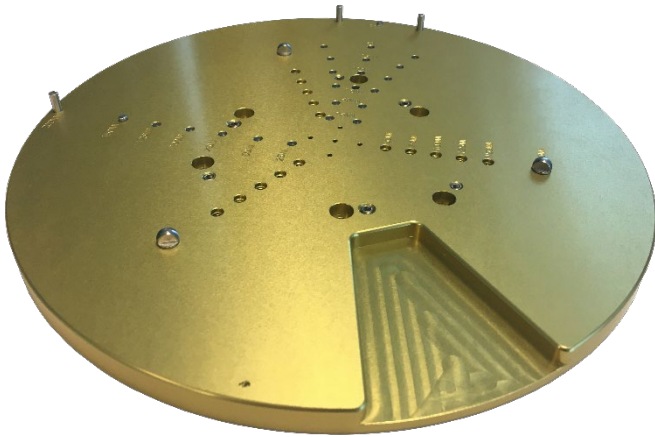


图 1. 用于应力测量的通用载台，带有三个定位销（在载台外缘）和三个支撑销钉（在载台中部）。定位销和支撑销钉的位置可调整，以适应不同的晶圆尺寸。

轮廓仪使用金刚石探针接触样品以收集形貌数据，沉积前后的数据收集采用相同的测试配方，以确保收集的数据具有相同的材料特性和数据分辨率。样品本征特性对于某些光学分析技术而言可能是个问题，而探针式轮廓仪的直接测量技术可以测量任何样品，不受样品特性的影响。应力测试配方的建立优化了数据采集，其关键参数包括扫描长度、探针施力、扫描速率、采样率和纵向量程。在应力测试配方中，用户还可以选择计算应力所用的算法工具。多项式拟合算法提供了最佳测量性，但也可以使用 13 点最小二乘拟合法来增强与先前应力测量技术的匹配度。



图 2. 用于应力测量的通用载台上的 200 毫米晶片。载台外围可以看到两个定位销，晶圆放置于三个支撑销钉（图中未显示）的顶部。

建议的测量参数包括：

- 探针针尖半径 $\geq 2\mu\text{m}$ ；
- 扫描长度应为晶圆总直径的 80%，扫描应通过晶圆中心；
- 探针施力为 2mg；
- 扫描速率为 1-5mm/s；
- 采样率为 200Hz；
- 针尖纵向量程取决于晶片的翘曲度。如果实现不知道翘曲度，请从最大的纵向量程开始测量，以适应较大的翘曲度。如果已知翘曲度跨越的针尖纵向运动范围较小，则使用较小的纵向量程。

应力测量准确性保障

Tencor 和 HRP 探针式轮廓仪采用光学平晶，以确保载台在水平面无起伏地横向移动。为确保轮廓仪针尖在曲面上运动的准确性，可测试已知曲率半径的镜面，如图 3 所示。表 1 列出了与扫描长度函数相关的平整度规格，以及翘曲度和应力重复性。

表 1. 扫描平整度和性规格。

	规格	参考曲面
扫描平整度*	30mm 扫描长度下为 40nm	1/20λ, 150mm 光学平面
	60mm 扫描长度下为 75nm	
	130mm 扫描长度下为 170nm	
翘曲度重复性*	0.1%	20m 半径镜面
应力重复性	2.5%	多项式拟合计算结果

*扫描平整度和翘曲度重复性基于 15 次重复次数、2mm/s 扫描速度、2mg 力、200Hz、最大针尖纵向范围、在 100mm 镜面上的 80mm 扫描长度和在 150mm 晶圆上的 120mm 扫描长度。

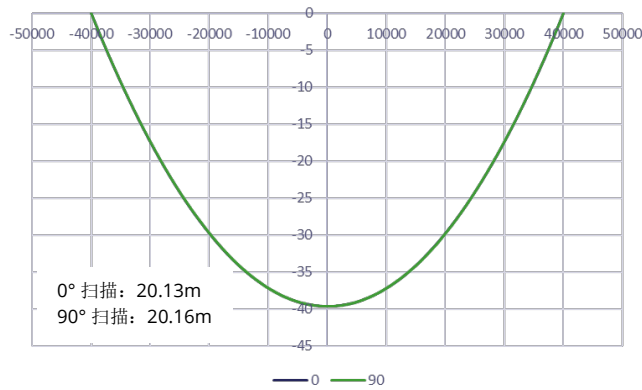


图 3. 标定半径为 $20.13 \pm 0.5\text{m}$ 的镜面测量结果。

应力测量分辨率

应力分辨率取决于三个参数：轮廓仪的纵向范围、衬底的弹性特性以及衬底和薄膜的厚度。对于 Tencor P 系列，针尖纵向范围取决于使用的扫描头（范围为 $6.5\mu\text{m}$ - $1000\mu\text{m}$ ）。而 HRP 的纵向在 $3.25\mu\text{m}$ 与 $327\mu\text{m}$ 之间可变。无论选择哪个范围，垂直分辨率都是亚埃级。衬底的双轴弹性模量因所用衬底而异。通常该值在 $1 \times 10^{11}\text{Pa}$ 与 $5 \times 10^{11}\text{Pa}$ 之间变化。最后，薄膜和衬底厚度对于确定应力分辨率来说非常重要。衬底厚度的典型值约为数百微米，而薄膜厚度通常在 100\AA 与 $2\mu\text{m}$ 之间，当然也可能有更厚的薄膜。表 2 列出了可以测得的应力值示例。

表 2. 一系列衬底的测定应力。

砷化镓	硅(100)	碳化硅
8MPa	132MPa	13.4GPa
8 英寸晶圆	8 英寸晶圆	6 英寸晶圆
$2\mu\text{m}$ 薄膜	500nm 薄膜	200nm 薄膜
$5\mu\text{m}$ 翘曲度	$25\mu\text{m}$ 翘曲度	$250\mu\text{m}$ 翘曲度

结论

HRP 系列和 Tencor P 系列轮廓仪可提供有效的薄膜应力监控和测量方法。由于过高应力可能会导致半导体器件失灵，因此应力监控至关重要。轮廓仪使用接触式量测方法，让探针针尖持续接触样品表面进行扫描，无需拼接即可进行大范围扫描，可确保数据准确性和重复性。这种直接测量不受样品材料和光学特性的影响，得到的结果具有亚埃级的垂直分辨率。应力配方中配有两种不同算法分析数据：13 点最小二乘拟合和高阶多项式拟合。根据薄膜沉积前后曲率半径的变化，使用 Stoney 方程式提供准确、可重复的薄膜应力测量结果。

参考文献

1. "Material Science of Thin Films", Chapter 12, p 711-712, Milton Ohring, 2nd Ed., 2002.
2. "Tencor P-11 Long Scan Profiler Operations", KLA, 1996, section 11.
3. Wafer Stress Application Option, Ch. 14.