arXiv:2302.00793v1 [physics.optics] 1 Feb 2023

用于向量-矩阵乘法的硅光子学平台上的逆向设计低折射率对比度结构

Vahid Nikkhah, [1] Ali Pirmoradi, [1] Farshid Ashtiani, [1]
Brian Edwards, [1] FiroozAflatouni, [1] Nader Engheta[1*]

[1]宾夕法尼亚大学电气与系统工程系，美国宾夕法尼亚州费城，邮编19104。

*通信作者；电子邮箱：engheta@seas.upenn.edu

逆向设计的硅光子超结构为利用电磁波执行模拟计算提供了高效平台。然而，由于计算上的困难，要将这类超结构扩展到处理大量数据通道并非易事。此外，典型的逆向设计流程采用较小的计算域，因此往往会利用谐振特征来实现设计目标。这会导致结构带宽窄，且对制造误差高度敏感。在此，我们采用一种基于有效折射率近似并施加低折射率对比度约束的二维逆向设计方法。由此得到的紧凑非晶透镜系统通常具有前馈特性，且谐振较弱。我们设计并通过实验演示了一个用于 2×2 与3×3矩阵的向量-矩阵乘积。我们还使用所提出的二维计算方法设计了一个 10×10矩阵。这些示例表明，这些技术具有

有望实现规模更大的基于波的模拟计算平台。

引言

能够执行数学运算和求解方程的光子结构正日益受到关注，这得益于光学模拟计算的复兴；光学模拟计算有望利用光实现低功耗、高速、并行计算（*1–11*）。尤其是，神经网络可以受益于向量-矩阵乘法：将其实现为光学模拟计算模块，可降低每次算术运算的能耗，同时显著提高速度（*12–16*）。从物理上实现用于向量-矩阵乘法的 $N \times N$ 矩阵，需要在单个器件中实现 N^2 个不同的目标。利用电磁波执行向量-矩阵乘法，可以表述为在k空间进行自由空间傅里叶光学（*1, 3, 17,18*），也可以将空间模式作为在实空间中运算的基底（*4,19*）。实现方式有两种。可以将 n 模式扩展为 n^2模式，利用额外维度；这些模式先被直接作用，再进行求和（*20*）。在三维器件中，可以采用基于玻恩近似定义的超表面来实现此类运算，并且只考虑其透射系数（*20,21*）。另一种方式是通过一系列混合运算对 n 模式进行操作，形成网格架构，通常只能实现“前向”运算（*22–25*）。最后，也可以利用逆向设计结构对模式进行运算，该结构可能包含内部共振（*1,4*）。

结合拓扑优化与形状优化，逆向设计*1,4*尤其适合用于设计此类结构。具体而言，基于密度的拓扑优化是一种将设计区域内的材料适当地离散为大量自由参数，且不预设结构形态的技术*26*。随后，利用伴随法等合适的数学方法优化这些参数；伴随法能够通过相对少量的

简单前向仿真*27,28*计算梯度。完成拓扑优化后，可将材料二值化，并矢量化为含有大量参数的形状，再使用伴随法对其进行类似的优化。

尽管硅光子学（SiPh）平台中的结构本质上是平面的，但其场并不具有二维对称性，因此设计人员往往需要采用三维仿真（*29*）。由于三维仿真的计算难度较高，而且达到最佳性能可能需要大量迭代，三维逆向设计仅适用于光学尺寸较小的结构（即数十个立方波长）。这实际上限制了可考虑的信息通道数量，因而也限制了矩阵规模。因此，需要一种高效的计算平台，以降低 SiPh 中全波三维仿真的计算成本。

因此，能够在大量输入和输出信息通道之间执行模拟计算的光子结构“降阶”近似至关重要。本文采用基于传播的二维有效折射率近似（p2DEIA）（*30–33*）来处理三维平面结构，从而显著降低计算量。通过设计硅厚度变化相对较小的 SiPh 结构，这些区域之间的有效折射率差也会较小。这种低折射率对比 可减少面外散射，提高 p2DEIA 近似的准确性，并减少反射（即减少共振，从而增加带宽）。我们通过设计和制造可执行向量-矩阵乘法的 SiPh 超构结构，验证了基于低折射率对比 p2DEIA 的逆向设计方法。

结果

问题描述

本文中，我们在 SiPh 平台上对给定矩阵M 执行向量-矩阵乘法，并使用一系列空间模式的复振幅表示向量。为

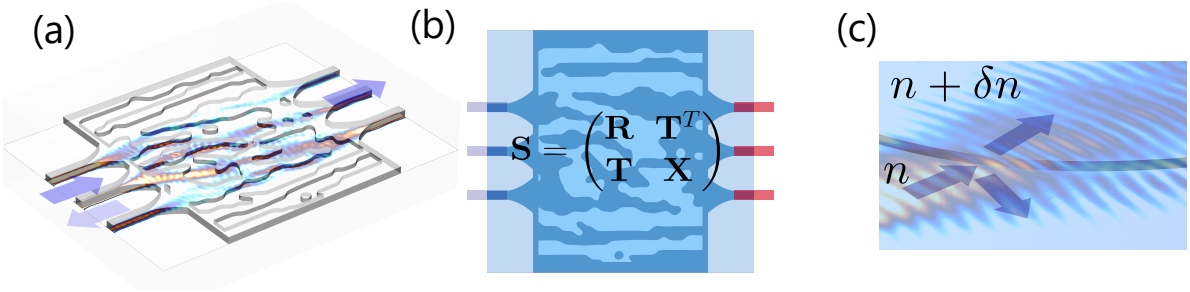

图 1：**硅光子平台中逆向设计结构的示意图**（a）由散射矩阵的透射部分组成的向量-矩阵乘法，S。（b）经过优化以实现该透射部分的通用结构。（c）低折射率对比带来的前馈特性：反射较小，主要表现为折射行为。

此目的，我们将单模硅波导阵列视为输入通道，另一个阵列则表示输出通道。输入和输出波导均连接至一个设计区域，其中硅厚度以参数化方式设定，在两个数值之间变化。（见图 1(a)）。一般而言，这类结构可以用散射矩阵完整描述；任意波导（指定为输入端或输出端）上的入射空间模式都会映射至所有其他模式（见图 1(b)）。通过指定输入端口注入的波可以用复值向量进行数学描述。虽然并非设计的重点，也可以类似地定义来自指定输出端口的入射向量。将入射波表示为两个串接的 $N \times 1$ 向量后，散射矩阵（$2N \times 2N$）可由四个分块矩阵（每个 $N \times N$）定义。由于我们关注的是从输入端口到输出端口的传播这一实现向量-矩阵乘积的物理过程，因此主要关注透射分块矩阵（$\mathbf{T} = \mathbf{M}$）。此外，我们还希望尽量减少反射回指定输入端口的波（即$\mathbf{R} = \mathbf{0}$）。因此，我们的目标是在二维设计区域内计算空间变化的硅厚度分布，以实现上述目标。

有效折射率近似

众所周知，三维结构的仿真与设计在计算上代价高昂，尤其是在任何迭代设计方法中。受限于硅薄板中的波以导波平板模式的形式沿面内几何结构传播。平板模的有效折射率取决于薄板的厚度，以及衬底、硅芯层和包层的折射率。例如，厚度为 $220\ nm$ 的硅层浸入折射率为 $n_{\mathrm{SiO_2}} = 1.44$ 的 $\mathrm{SiO_2}$ 中，且具有 $n_{\mathrm{Si}} = 3.48$，支持一种基模平板模式，其 $n_{\mathrm{eff,\ 220nm}} = 2.86$ 在 $\lambda_0 = 1.525\ \mu m$ 处。硅芯层厚度减小时，有效折射率也会降低。例如，对于厚度以类似方式定义为 $150\ nm$ 的硅层，基模平板模式的有效折射率降至 $n_{\mathrm{eff,\ 150nm}} = 2.56$。在一定约束条件下，可以构建二维计算模型，通过使用平板模的有效折射率来表示不同的硅芯层厚度，从而近似模拟波在三维平面结构中的面内传播（即导波平板模式）。例如，二维模型中折射率为$n = n_{\mathrm{eff,\ 220nm}} = 2.86$ 的区域，能够完整、准确地模拟硅厚度为 $220\ nm$ 的三维结构中基模平板模式的面内传播。本文将此模型称为基于传播的二维有效折射率近似（p2DEIA）。

假设有两个区域，一个区域的硅层厚度为 220nm，另一个区域经过均匀刻蚀，厚度更小（见图1(c)）。当基模平板模式到达这两个区域之间的界面时，可能发生多种现象。p2DEIA 模型将通过斯涅尔定律有效地捕捉基模平板模式的反射角和透射角（见图1(c)）。然而，三维高度不连续性还会产生其他效应。例如，部分波会从硅薄板散射出去，进入 $\mathrm{SiO_2}$ 包层。此外，硅芯层内还可能激发高阶平板模式。由于存在这些额外的散射通道，三维高度不连续性只能 用二维模型近似。二维模型的精度取决于

与这些其他散射通道的耦合强度。当硅厚度相近时，基模平板模式之间的模态重叠非常大，因此，所有其他耦合都必定很弱（见补充材料图1）。因此，对于足够浅的刻蚀深度，可以假设与未建模开放通道的耦合可以忽略，并放心使用近似的 p2DEIA方法，即使在界面处也能获得较高精度。

基于前文定义的基模平板模式折射率的 p2DEIA 方法，能够准确描述波在均匀刻蚀区域内的传播、折射角和反射角。然而，与其他模型相比，它对界面处反射和透射的幅度及相位描述可能不够准确。例如，在基于变分方法的模型中，人们会选择一个厚度，并预期波的大部分能量会分布在该厚度范围内（*29*）。该区域的有效折射率按与上述类似的方法计算，能够准确模拟传播。然而，其他所有有效折射率都以所选区域为参照进行定义，旨在准确模拟界面行为，但代价是无法准确模拟其他区域内的传播（*29*）。对于基于全内反射（TIR）和光栅耦合器的波导结构，这些变分方法可能更为合适，因为准确模拟区域之间的界面至关重要。然而，如果能量在较长距离内大量传播穿过刻蚀区域，这些方法的准确性就会较低，因为受影响折射率的折衷会导致相位误差不断累积。我们将其称为基于界面的有效折射率近似（i2DEIA）。

如前所述，p2DEIA 的准确性要求界面两侧的硅厚度相对差异不大，因此我们将仅考虑满足这一条件的设计。因此，各区域对应的有效折射率将具有低折射率对比，使界面反射幅度较小、透射幅度较大。因此，可以预期，波主要从输入端口向输出端口传播，并在从一个区域进入另一个区域时于界面处发生衍射。缺乏显著的

反射将形成一种低共振前馈光学系统，其表现类似一系列形状不规则的透镜。从实验角度看，这类系统自然具有宽带光学响应，并且只要界面角度总体保持不变，对界面形状的微小扰动就不太敏感。这些特性十分实用，因为人们非常希望超结构能够抵御诸如“过度蚀刻”之类的制造误差。另一方面，低折射率对比界面只能产生较弱的衍射，与高折射率对比界面相比，对波操控的控制能力也较弱。

无源性对几何定义的影响

任何由边界 S 包围的光子结构，只要仅表现出无损传播，就可以用酉散射矩阵来描述。可以对这类边界上的模式进行编号。其中一些模式将被指定为输入（$\mathbf{I}$）、输出（$\mathbf{O}$），其余模式则被简单地视为“吸收”（$\mathbf{A}$）。在所有情况下，这些模式上的任何出射波（$\mathbf{I}_{\text{out}}$、$\mathbf{O}_{\text{out}}$、$\mathbf{A}_{\text{out}}$）都不会返回系统。除了分块矩阵$\mathbf{T}$ 和$\mathbf{R}$之外，加入向量$\mathbf{A}$ 也使我们能够将其他分块矩阵纳入概念框架（χ_1等）。

$$\begin{pmatrix} \mathbf{I}_{\text{out}} \\ \mathbf{O}_{\text{out}} \\ \mathbf{A}_{\text{out}} \end{pmatrix} = \mathbf{S} \begin{pmatrix} \mathbf{I}_{\text{in}} \\ \mathbf{O}_{\text{in}} \\ \mathbf{A}_{\text{in}} \end{pmatrix} \qquad \mathbf{S} = \begin{pmatrix} \mathbf{R} & \mathbf{T}^T & \chi_{\mathbf{1}} \\ \mathbf{T} & \mathbf{X} & \chi_{\mathbf{2}} \\ \chi_{\mathbf{3}} & \chi_{\mathbf{4}} & \chi_{\mathbf{5}} \end{pmatrix} \tag{1}$$

这些吸收模式永远不会受到照射（即$\mathbf{A}_{\text{in}} = \mathbf{0}$），而且不会对这些模式的精确形状进行表征。不过，这些模式确实会影响系统。

由于散射矩阵是酉矩阵，$\mathbf{S}$ 的所有奇异值均为 1。作为酉矩阵中的一个分块矩阵，$\mathbf{T}$ 的奇异值必须小于 1。此条件称为无源性。当不受约束的吸收模式数量非常大时，$\mathbf{S}$ 为酉矩阵这一条件不再构成显著约束，而对$\mathbf{T}$ 施加无源性条件便已足够。另一方面，如果吸收模式的数量很少（即只有一个），则无法保证

即使散射矩阵由满足无源性的$\mathbf{T}$、$\mathbf{R} = \mathbf{0}$ 和不受约束的$\mathbf{X}$ 定义，它也属于酉矩阵空间。因此，在本研究中，我们假设不受约束的$\mathbf{A}$代表数量足够多的吸收模式。在这种情况下，我们可以利用无源性条件来选择目标透射矩阵。在由透射矩阵上的无源性约束所描述的向量空间中，我们任意选取以下矩阵，用于一个 2×2 和 3×3 光网络：

$$\mathbf{T}_{2\times2} = \begin{pmatrix} 0.43 + j0.43 & -0.47 + j0.12 \\ 0.43 + j0.22 & 0.51 - j0.33 \end{pmatrix} \qquad \mathbf{R}_{2\times2} = \mathbf{0} \tag{2}$$

$$\mathbf{T}_{3\times3} = \begin{pmatrix} 0.49 + j0.24 & -0.53 + j0.44 & 0.44 - j0.16 \\ 0.60 + j0.41 & 0.20 - j0.46 & -0.32 - j0.36 \\ 0.41 + j0.11 & 0.5 + j0.15 & 0.20 + j0.71 \end{pmatrix} \qquad \mathbf{R}_{3\times3} = \mathbf{0} \tag{3}$$

超构结构配置与逆向设计

为实现目标透射矩阵，所提出的光学超构结构基于硅光子学平台在 $\lambda_0 = 1.525\ \mu m$上进行设计。采用 p2DEIA 模型设计的优化3×3 和 2×2 结构的二维示意图分别见图 2(a) 和 (d)。每条硅波导的宽度为 $500\,nm$，厚度为 $220\,nm$。对于 2×2 情形，设计区域由图 2(a) 中黑色虚线边界内的区域界定，是一个宽度为 $11\,\mu m$、长度为 $10.3\,\mu m$的矩形区域。对于 3×3情形，该区域为椭圆，宽度为 $34\,\mu m$，长度为 $23\,\mu m$ （见图 2(d)）。在这两种结构中，红色虚线所示的吸收体边界沿两侧延伸多个波长，从而提供足够多的吸收模式$\mathbf{A}$，这是设计过程所必需的。

优化问题在数学上可描述如下

$$\begin{aligned} &\min_{n(x,y)} \|\mathbf{T} - \mathbf{T}_{\text{targ}}\|_F + \|\mathbf{R} - \mathbf{0}\|_F \\ &\text{s.t. } n_{\text{eff, 150nm}} \le n(x,y) \le n_{\text{eff, 220nm}} \end{aligned} \tag{4}$$

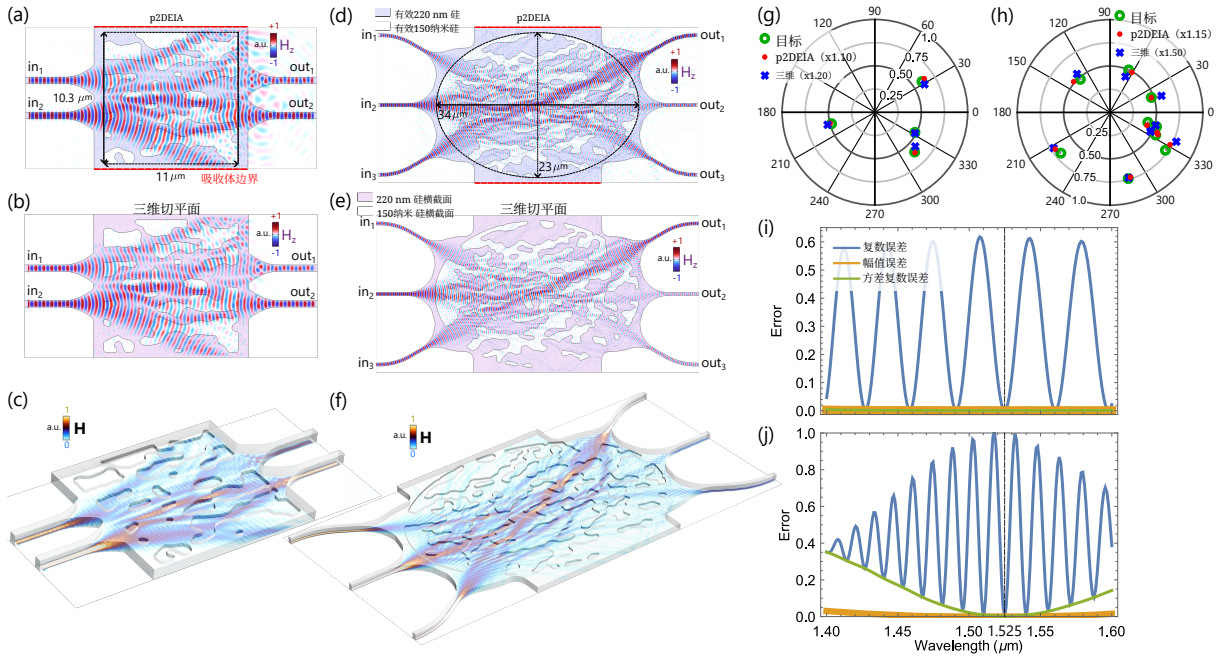

图 2：**基于硅光子学、用于执行 2×2 和 3×3 矢量-矩阵乘法的逆向设计超结构**。设计与仿真在自由空间波长 $\lambda_0 = 1.525\,\mu m$下进行。(a) 2×2 和(d) 3×3 示例的 p2DEIA 优化结构，以及任意输入矢量激励下的面外磁场分布时间快照。(b)和(e) 与 p2DEIA 中相同激励下，全波三维渲染结构中穿过硅的切平面上的磁场分布时间快照。为便于与 p2DEIA 情况（即(a)和(d)）的结果比较，此处采用垂直视角。(c)和(f) 设计结构三维视图中穿过硅的磁场分布绝对值。(g)和(h) 2×2 和 3×3 结构的目标传输矩阵元（绿色圆圈）、p2DEIA 值（经适当调整因子校正，括号内所示）（红色圆圈）以及三维值（经适当调整因子校正，括号内所示）（蓝色圆圈）。(i)和(j) 2×2 和 3×3结构中，目标传输值与三维传输值之差的平方的归一化弗罗贝尼乌斯范数随波长的变化（即 $\|\mathbf{T} - \mathbf{T}_{\text{targ}}\|_F^2/4N$）。

其中，$\|\cdot\|_F$ 为弗罗贝尼乌斯范数。逆向设计所定义的标量目标函数表示复值目标传输矩阵与当前迭代中由结构实现的传输矩阵之间的“距离”。目标函数中还加入了用于最小化反射的类似表达式。使用 p2DEIA 时，设计区域由有效折射率分布参数化，其范围限定在 $n_{\text{eff, 150nm}}$ 和 $n_{\text{eff, 220nm}}$之间。通过在 COMSOL Multiphysics® (*34*) 中采用基于伴随法的基于密度的拓扑优化逆向设计方法，对 p2DEIA 模型中的有效折射率分布进行优化，以最小化标量目标函数。

在逆向设计过程中，使用一个逐渐变陡的 S形投影函数，使有效折射率分布趋向于极值 $n_{\text{eff, 150nm}}$ 和 $n_{\text{eff, 220nm}}$ 。随后，将优化分布矢量化，以定义与 $n_{\text{eff, 220nm}}$ 和 $n_{\text{eff, 150nm}}$相关的两类区域，分别如图 2(a) 和 (d) 设计区域内高亮显示和未高亮显示的互补区域所示。这些区域最终确定了三维结构中用于仿真和制造的相应厚度，如图 2(c) 和 (f) 所示。

p2DEIA 模型是一种计算效率较高的近似模型，其性能依据与该结构完整三维仿真结果的吻合程度进行评估；本文将完整三维仿真结果称为真实值。图 2(a) 和 (d) 显示了矢量化 p2DEIA 模型中面外磁场分布随时间变化的快照；图 2(b) 和 (e) 则显示了在与 p2DEIA 模型相同的激励下，三维渲染结构中穿过硅薄板的切面上的相应场分布。将 p2DEIA 模型的仿真场分布与三维结构进行视觉比较，可以看到二者具有相似特征，表明 p2DEIA 能够成功近似光在三维结构中的面内波传播。最后，图 2(c) 和 (f) 展示了所设计超构结构的三维视图中，穿过硅的切平面上的磁场分布绝对值。

在图 2(g) 和 (h) 中，我们比较了目标传输矩阵（$\mathbf{T}^{\text{targ}}$）、p2DEIA 模型实现的矩阵（$\mathbf{T}^{\text{p2DEIA}}$）以及三维模型实现的矩阵（$\mathbf{T}_{3\text{D}}$）的复数值。我们允许通过归一化消除两类误差。首先，对$\mathbf{T}^{\text{p2DEIA}}$ 中的所有值进行恒定幅度调整，以补偿因面外散射造成的未建模能量损耗。其次，对每类波导分别施加恒定相位调整，并应用于$\mathbf{T}^{\text{p2DEIA}}$中的所有值。这是因为与 i2DEIA 模型相比，p2DEIA 在对波导进行建模时效果较差。换言之，波导中的导模具有略微不同的波数，因此 p2DEIA 与

i2DEIA 和三维模型之间会产生相位误差。完成这些简单的调整后，可以看到目标值、p2DEIA 值和三维模型值高度吻合，这表明我们成功实现了所需的透射系数，同时也证明 p2DEIA 在近似三维结构方面具有较高精度。

接下来，在图 (i) 和 (j) 中，我们通过考察三种误差来分析所实现三维设计的光谱响应。在所有情况下，我们都用 $4N$ 对平方误差进行归一化，这样一来，大小为 N 、具有 180 度相位误差的酉矩阵对应 100%。参见补充材料。橙色曲线表示透射参数绝对值之间的弗罗贝尼乌斯距离。$(\||\mathbf{T}_{3\mathrm{D}}(\lambda)| - |\mathbf{T}_{\mathrm{targ}}|\|_F^2/4N)$。目标值与模拟值之间的距离很小，表明随着波长变化，透射参数的振幅基本保持不变，并且接近目标值。然而，对于旨在执行复数 矩阵乘法的器件而言，这并不足以作为成功与否的判据。蓝色曲线表示不同波长下的透射参数与目标值 $\|\mathbf{T}_{3\mathrm{D}}(\lambda) - \mathbf{T}_{\mathrm{targ}}\|_F^2/4N$之间的弗罗贝尼乌斯距离。在设计波长处，该距离相当小，表明在应用上述振幅和相位校正后，三维结构能够以很高的精度实现目标透射参数。然而，误差会随波长变化而呈现振荡行为。这是由于输入端口与输出端口之间的传播时延会在透射参数中引入全局相位误差。从概念上说，这与光沿正向通过任意介质传播时，波长改变所引起的相位旋转有关（即 $e^{j(k(\lambda)-k(\lambda_{0_{))}}l)}$ ）。就相位差而言，这一全局相位无关紧要，而通常真正重要的是这些相位差。因此，我们考虑一种误差：对所有值应用全局相位校正（绿色曲线） $\phi(\lambda)$，使得 $\mathbf{T}'_{3\mathrm{D}}(\lambda) = \mathbf{T}_{3\mathrm{D}}(\lambda)e^{-j\phi(\lambda)}$ ，从而最小化 $\|\mathbf{T}'_{3\mathrm{D}}(\lambda) - \mathbf{T}_{\mathrm{targ}}\|_F^2/4N$。正如预期，这消除了误差的振荡特征，因此我们可以得出结论：所设计的超结构能够执行向量-矩阵

运算，并且适用于较宽的波长范围，因此可作为用于此类模拟计算的宽带光学结构。

实验结果

设计的超构结构已完成制造，实验结果如图3所示。在图3(a)和(c)中，分别对比了 2×2 和 3×3 结构的目标透射率参数、p2DEIA中的模拟透射率、三维模型中的模拟值，以及在 λ_0 =1.525 μm 处测得的透射率。如前所述，由于p2DEIA未对面外散射进行建模，三维结构中的少量能量会从硅薄板散射出去，因此透射率略有降低。因此，为了直观地将三维结构的透射率与p2DEIA的透射率进行比较，将 2×2 的三维值乘以1.2，将 3×3 的三维值乘以1.5。应用调整因子后，目标值、p2DEIA值和三维值彼此吻合。为了直观地将测得的透射率与目标值进行比较， 3×3 和 2×2 分别需要采用类似的比例因子1.45和1。2。因此，测得的透射率与三维模型的数值十分接近，表明该模型具有很高的保真度。总之，尽管p2DEIA模型存在限制，测量也并非完美，但在 $\lambda_0 = 1.525\ \mu m$ 处，目标值、p2DEIA值、三维值与测得的传输值吻合良好，这表明p2DEIA作为一种用于三维平面结构逆向设计的计算工具是成功的。

在图 (b) 和 (d) 中，展示了透射率的波长扫描结果。粗实线表示来自三维模型的仿真数据，细实线表示测量数据。在设计波长处，仿真数据与测量数据吻合极佳（测量详情见补充材料）。如前所述，从.....可以看出，透射参数的幅值呈现出宽带响应。

仿真数据和实验数据随波长缓慢变化。

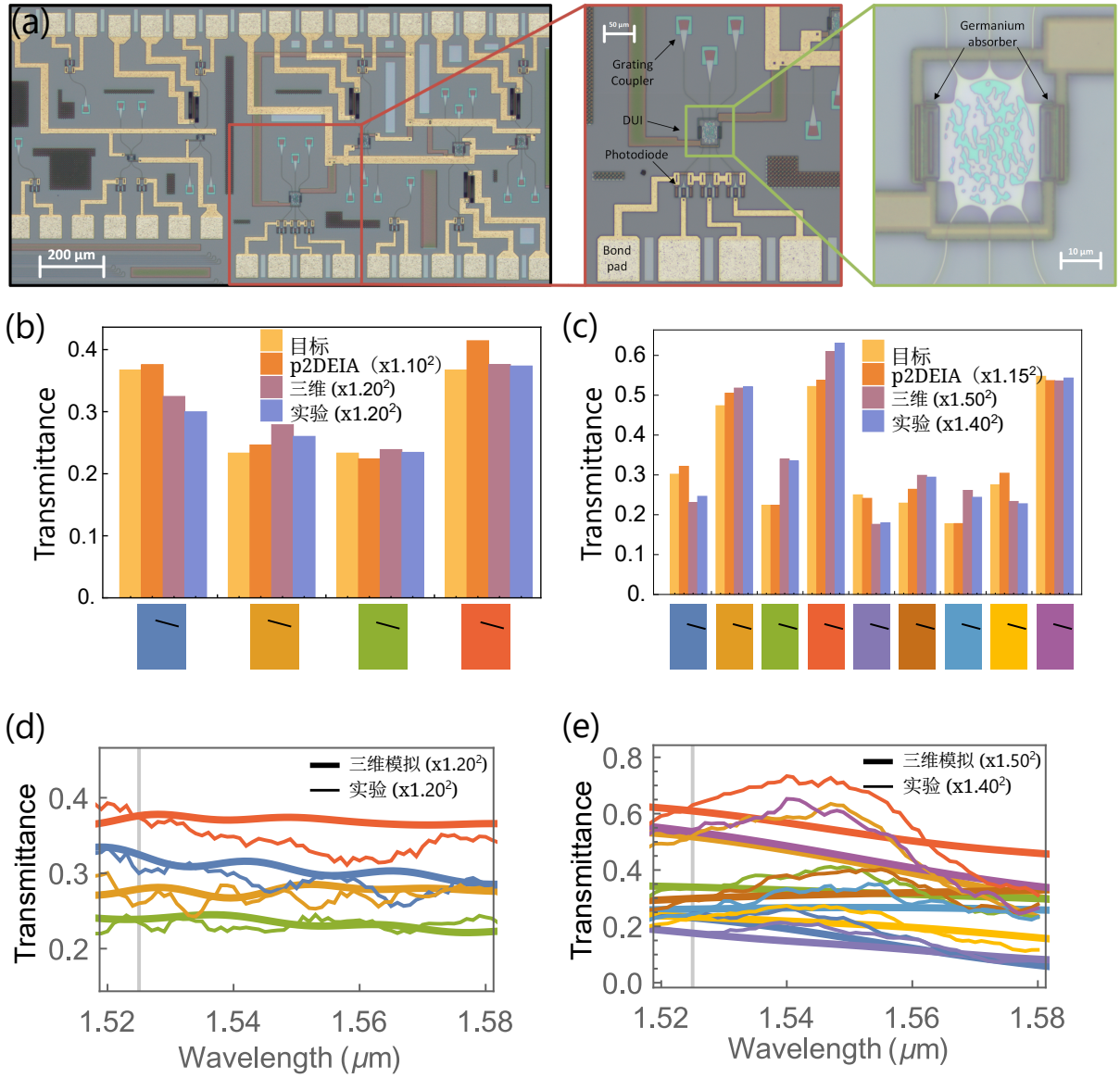

图3：仿真结果与实验结果的比较。(a) 实验显微照片，展示了若干 2×2 和 3×3 卷积核以及各种校准结构。突出显示了 3×3 卷积核。(b) 和 (c) 分别比较了 $\lambda_0 = 1.525\ \mu m$ 处 2×2 和 3×3 超构结构的测量透射率值、p2DEIA 和三维模型中的仿真透射率值以及目标值。(d) 和 (e) 分别展示了2×2 和 3×3 超构结构的测量透射率值（粗实线）和三维仿真结果（细实线）随波长的变化。

用于 10×10 向量-矩阵乘法的硅光子学超结构逆向设计

如前所述，p2DEIA 能够以可承受的计算工作量对光学大尺寸结构进行仿真和设计。这使得我们可以通过增加波导数量来扩展输入和输出数据通道。为此，我们采用相同的逆向设计技术，利用 p2DEIA 探索并设计一种能够执行 10×10 向量-矩阵乘法的结构（见图4(a)）。该结构在 $35.4\ \mu m \times 29.4\ \mu m$ 设计区域内设计，采用与前述示例相同的硅光子学平台和波长。设计区域跨越多个导波波长，即 $66\lambda_g \times 55\lambda_g$ 其中 $\lambda_g = \lambda_0 / n_{\text{eff, 220nm}}$。目标矩阵$\mathbf{M}_{10\times10}$ 在无源传输矩阵空间内任意选取（见 SM）。

图4(a)展示了优化后的几何结构，以及在 p2DEIA 模型中施加任意矢量激励时，面外磁场分布的一个时间截面。图4(b)展示了相同激励通过三维渲染结构时磁场分布的绝对值。对于如此巨大且光学尺寸极大的结构，无论采用频域还是时域仿真，进行全波三维仿真都十分困难；然而，借助 p2DEIA，我们能够在逆向设计过程中对该结构进行数百次迭代仿真。图4(c)展示了目标传输矩阵元素的复数值、优化后的 p2DEIA 模型所得数值，以及三维渲染结构所得数值。与 2×2 和 3×3 结果类似，p2DEIA 和三维仿真所得数值经过适当的调整因子校正后彼此对齐。为便于比较，相应的传输值以复平面极坐标图的形式表示，如图4(d)所示。（相应的调整因子标注在括号中。）

目标值与 p2DEIA 数值高度吻合，表明逆向设计顺利收敛。尽管如此，三维数值与 p2DEIA

数值仍相当吻合，这得益于采用了较粗网格和降低阶数的吸收边界条件；由于三维仿真的规模庞大，这些做法是不得已而为之。不过，p2DEIA 模型在每个界面都会引入一些误差（见补充材料）。更大的结构可能需要更复杂的界面建模，并应考虑面外散射，以及非传播模态中的储存能量所导致的相位不连续。尽管如此，从目标值、p2DEIA 和三维透射参数之间的高度吻合来看，优化设计表现良好。因此，该设计证明 p2DEIA 成功设计出一种光学大尺寸结构，可用于操控大量输入和输出数据通道。

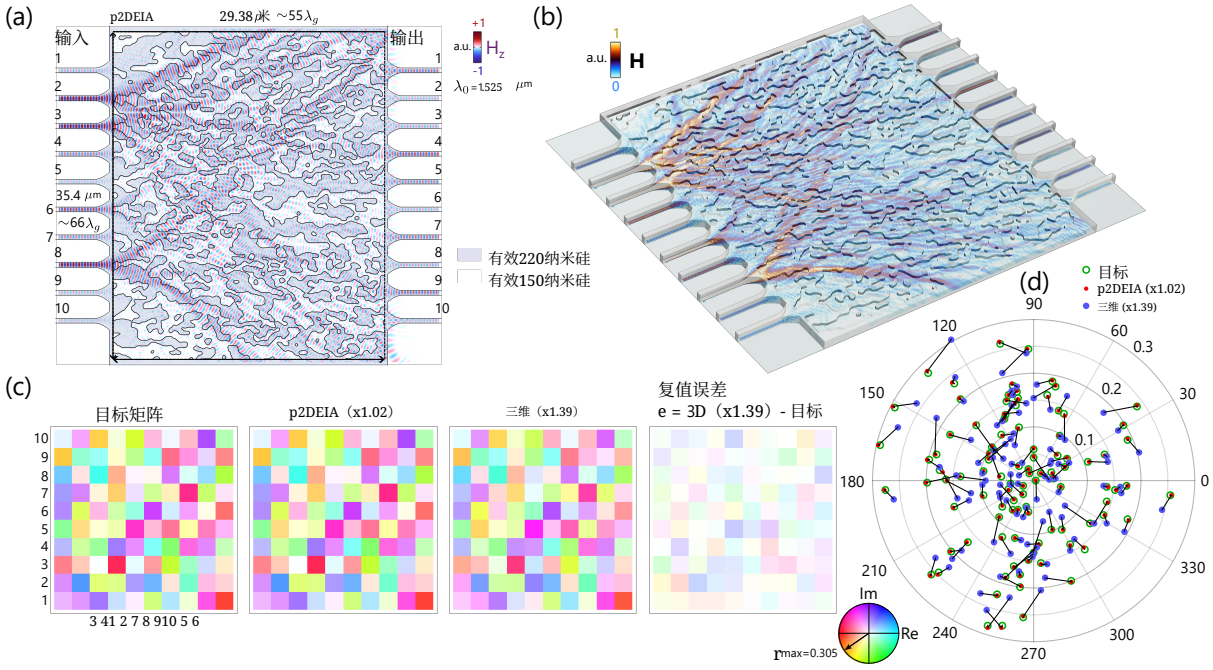

图4：**用于实现** 10×10 **向量-矩阵乘法的硅光子学超构结构逆向设计**(a) p2DEIA中的优化几何结构，以及任意输入激励向量下的面外偏振磁场分布。(b) 对于与(a)中相同的激励，三维渲染结构中硅切面上的磁场分布绝对值。(c) 目标传输矩阵、p2DEIA和三维模型中的模拟透射参数，以及目标传输参数与三维透射参数之间的复值误差。(d) 复平面极坐标图中对应的传输值。

总结与结论

本文采用 p2DEIA，并以平板模的有效折射率作为高效计算平台，对硅光子学中的三维平面结构进行逆向设计。假设低折射率对比，我们展示了经过逆向设计、可执行2×2和 3×3 向量-矩阵乘积的超构结构。仿真和实验结果表明，p2DIEA 在近似三维平面结构方面具有很高的精度。此外，p2DEIA 所依赖的低折射率对比这一关键假设，可实现前馈、低谐振结构，其响应随波长缓慢变化。因此，使用 p2DEIA 得到的结构在执行模拟计算时具有宽带光学响应，而且对制造误差也不那么敏感。由于 p2DEIA 是三维结构的降维近似，与全波三维仿真相比，将其扩展至更大结构所需的计算量要小得多，因此能够对用于大规模向量-矩阵乘法的超构结构进行逆向设计。我们通过设计一个可执行 10×10 向量-矩阵乘积的结构验证了这一点。

致谢

本研究部分由美国空军科学研究办公室（AFOSR）多学科大学研究计划（MURI）资助，项目编号为 FA9550-21-1-0312（资助对象：N.E.）；另有部分由美国海军研究办公室（ONR）资助，项目编号为 N00014-19-1-2248（资助对象：F.A.）。

参考文献

1. A. Cordaro,等，《自然·纳米技术》，*2023*年*1*月*12*日 （2023）。

2. V. Nikkhah, D. C. Tzarouchis, A. Hoorfar, N. Engheta，《ACS 光子学》（2022年）。

3. A. Silva 等，《科学》343，160（2014年）。4. N. Mohammadi Estakhri, B. Edwards, N. Engheta，《科学》363，1333（2019年）。5. H. Zhang 等，《自然·通讯》12，1（2021年）。6. M. Camacho, B. Edwards, N. Engheta，《自然·通讯》12，1（2021年）。7. W. Fu 等，《光：科学与应用》11，1（2022年）。8. D. Pan 等，《光子学研究》9，1758（2021年）。9. F. Zangeneh-Nejad, D. L. Sounas, A. Al`u, R. Fleury，《自然·材料评论》6，207（2021年）。10. T. Zhu 等，《自然·通讯》8，1（2017年）。11. A. Pors, M. G. Nielsen, S. I. Bozhevolnyi，《纳米快报》15，791（2015年）。12. T. Wang 等，《自然·通讯》13，1（2022年）。13. B. J. Shastri 等，《自然·光子学》15，102（2021年）。14. Y. Shen 等，《自然·光子学》11，441（2017年）。15. X. Lin 等，《科学》361，1004（2018年）。16. J. Feldmann 等，《自然》589，52（2021年）。17. H. Wang, C. Guo, Z. Zhao, S. Fan，《ACS 光子学》7，338（2020年）。18. T. Zhu 等，《自然·通讯》12，1（2021年）。19. H. Goh, A. Al`u，《物理评论快报》128，073201（2022年）。

20. N. H. Farhat, 《IEEE Expert》*IEEE Expert* **7**, 63（1992）。21. G. Wetzstein, 等, 《自然》*Nature* **588**, 39（2020）。22. D. A. Miller, 《光子学研究》*Photonics Research* **1**, 1（2013）。23. W. Bogaerts, 等, 《自然》*Nature* **586**, 207（2020）。24. W. R. Clements, P. C. Humphreys, B. J. Metcalf, W. S. Kolthammer, I. A. Walmsley, 《光学》*Optica***3**, 1460（2016）。25. D. C. Tzarouchis, M. J. Mencagli, B. Edwards, N. Engheta, 《光：科学与应用》*Light: Science & Applications***11**, 1（2022）。26. M. P. Bendsoe, O. Sigmund, 《拓扑优化：理论、方法与应用》（施普林格科学与商业媒体，2003）。27. S. Molesky, 等, 《自然·光子学》*Nature Photonics* **12**, 659（2018）。28. T. W. Hughes, M. Minkov, I. A. Williamson, S. Fan, 《ACS 光子学》*ACS Photonics***5**, 4781（2018）。29. M. Hammer, O. V. Ivanova, 《光学与量子电子学》*Optical and quantum electronics* **41**, 267（2009）。30. R. Knox, P. Toulios, 《亚毫米波研讨会论文集》（纽约布鲁克林, 1970），第20卷，第497–515页。31. K. S. Chiang, 《应用光学》*Applied optics* **25**, 348（1986）。32. G. B. Hocker, W. K. Burns, 《应用光学》*Applied optics* **16**, 113（1977）。33. K. Van De Velde, H. Thienpont, R. Van Geen, 《光波技术杂志》*Journal of Lightwave technology* **6**, 1153（1988）。34. Comsol 多物理场® v. 5.6。www.comsol.com。comsol ab，瑞典斯德哥尔摩（2020）。

35. Y. Saad，M. H. Schultz，*SIAM* 科学与统计计算杂志**7**，856（1986）。

36. O. Schenk, K. G¨artner，《并行计算》**28**，187（2002年）。

补充材料

有效折射率近似概述

人们对基于波的模拟计算日益浓厚的兴趣，对能够高效仿真和设计光学尺寸较大的电磁结构的数值方法提出了需求。缓解计算负担的一种方法，是用有效模型近似某个结构。有效模型旨在用一个更简单的问题替代原始问题，同时保留相关电磁特征的本质特征。原始问题的计算量可能会随着结构尺寸的增大而快速扩展；然而，其有效模型的计算量相对于计算资源而言，可能以更有利的方式扩展。构建有效模型的一种方法是进行降维。一个例子是本征模态展开法（EME）：将具有任意横截面（但在 z 中均匀）的波导约化为若干导模，每个导模都根据其折射率进行一维传播。这消除了两维计算负担，而剩余的一维则可通过时域有限差分法（FDTD）等适当方法处理。此外，在平面介电结构的变分法中，可以用二维有效折射率分布替代三维结构，以模拟实际三维结构中的面内波传播（*29*）。通过最小化麦克斯韦方程组的泛函形式，可以优化有效折射率分布，从而估算面内波传播。

在本研究中，我们采用基于平板模有效折射率的传统有效模型（p2DEIA），对基于硅光子学的三维平面结构进行逆向设计。在典型的硅平面结构中，如图 S1(a) 所示，波被限制在硅芯层

内，以平板模的形式传播。通过在所需区域向下刻蚀以改变芯层厚度，可以改变平板模的传播特性。刻蚀区域与未刻蚀区域之间形成的不连续性会产生界面，波在这些界面处发生反射和折射；通过设计界面的拓扑和形状，可以操控从一侧传播到另一侧的光。对麦克斯韦方程组进行完整的三维全波仿真，计算成本可能会使设计过程极为低效。具体而言，在某些方法（例如逆向设计）中，计算成本会随着仿真次数增加而增长，因为优化过程中需要对结构进行数百次仿真。另一方面，在硅平面结构中，波在不同区域内主要以基模平板模式传播（见图 S1(a)）。因此，只要波在通过刻蚀区域与未刻蚀区域之间的界面时耦合至基模平板模式，其沿面外方向（图 S1(a) 中的 y-方向）的波函数就是已知的；也就是说，各区域中的基模平板模式与高阶模之间的耦合可以忽略不计。因此，为了高效仿真波在图 S1(a) 所示平面结构中的传播，可以利用面内平板模式传播（图 S1(a) 中的 $xz-plane$），采用有效二维模型进行仿真。如图 S1(b) 所示，在有效二维模型中，图 S1(a) 所示结构的区域和界面被映射到二维平面上。因此，满足亥姆霍兹波动方程的二维标量波代表平板模在整个结构中的面内传播。二维模型中各区域的折射率取自三维结构中相应区域所限制的平板模有效折射率。通过这种建模，只要平板模传播时未遇到界面，就能准确模拟其沿面内几何结构的相位演进。换言之，利用 p2DEIA 可以准确模拟体相位演进。着眼于

相位演进的精度会受到影响，因为遇到了界面 e

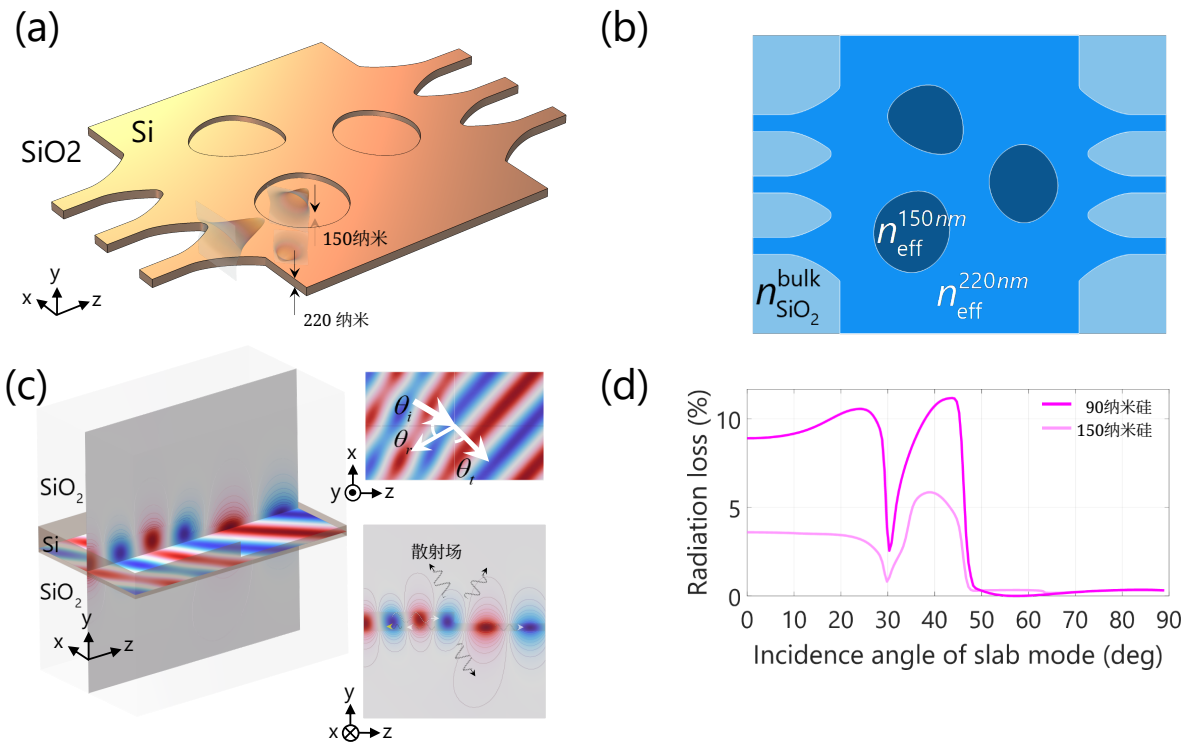

图 S1：**平面结构的有效折射率近似**(a) 具有刻蚀区域、用于调控面内波传播的典型硅平面结构三维模型；(b) 将三维模型投影到有效二维模型（p2DEIA）的有效折射率近似方法，用于捕捉面内波传播的主要特征；(c) 平板模式传播以入射角 θ_i 入射到刻蚀界面时的全波三维仿真。在 xz 平面上，展示了由 p2DEIA 捕捉的主要面内波传播。在 yz 平面上，展示了由高度不连续性引起的次级面外散射，这是 p2DEIA 无法捕捉的；(d) 面外散射（辐射损耗）的能量随入射角 θ_i 的变化。

。如图 S1(c) 所示，界面行为的真实值是完整的三维模型。波在三维界面中的传播会产生二维模型未涵盖的一些效应，例如图 S1(c) 中 yz 平面上的场分布所示的面外散射。相比之下，使用 p2DEIA 模型模拟平面波照射两种有效折射率不同材料之间的平面界面，将得到标准的菲涅耳反射与透射系数。为了深入理解 p2DEIA 模型引入的界面误差，我们可以将菲涅耳反射与透射系数与基模平板模式的反射和透射进行比较

在完整的三维模型中，作为入射角的函数。

向次级模式的耦合（面外散射），即我们所称的辐射损耗，会使基模平板模式的反射和透射略微偏离使用二维模型计算得到的菲涅耳反射与透射系数。在图 S1(d) 中，我们绘制了耦合至辐射损耗的输入功率占比（即$L = 1 - |R|^2 - |T|^2$），随入射角变化的情况，分别针对 $220\ nm$ 至 $150\ nm$ 界面和 $220\ nm$ 至$90\ nm$ 界面。

可以看出，辐射损耗的大小与刻蚀高度成正比，同时呈现相似的角度色散。此外，辐射损耗在入射临界角附近达到峰值；超过该角度后会发生全内反射，辐射损耗随之减小。因此，p2DEIA 的准确性与刻蚀高度有关：刻蚀高度越大，p2DEIA 越不准确。在我们的设计中，我们采用相对较浅的刻蚀 $70\ nm$ （对应于厚度为 $150\ nm$的硅薄板），以控制 p2DEIA 的误差。请注意，尽管 $150\ nm$ 在临界角之前达到最大辐射损耗 $\approx$ % 6，但考虑所有可能角度时，平均辐射损耗更低。

图 S2 绘制了不同刻蚀深度界面处透射系数的振幅和相位。可以看出，在临界角之外，三维模型与 p2DEIA 的数据相吻合，这是因为全内反射导致的辐射损耗可忽略不计。在 p2DEIA 模型中，130nm 刻蚀的有效折射率差大于 70nm 刻蚀的有效折射率差，因此其透射幅值自然会降低。然而，这一差异很小，且两者的透射率都接近于 1。在临界角之前，由于面外散射模式，与 70nm 刻蚀相比，130nm 刻蚀的三维模型与 p2DEIA 之间存在明显更大的差异。透射率下降的原因在于基模之间的模式重叠（包括形状和模式中心）较差，从而导致反射增加、三维辐射损耗以及散射进入高阶导模（比较 p2DEIA 与三维模型反射的结果见图 S3）。

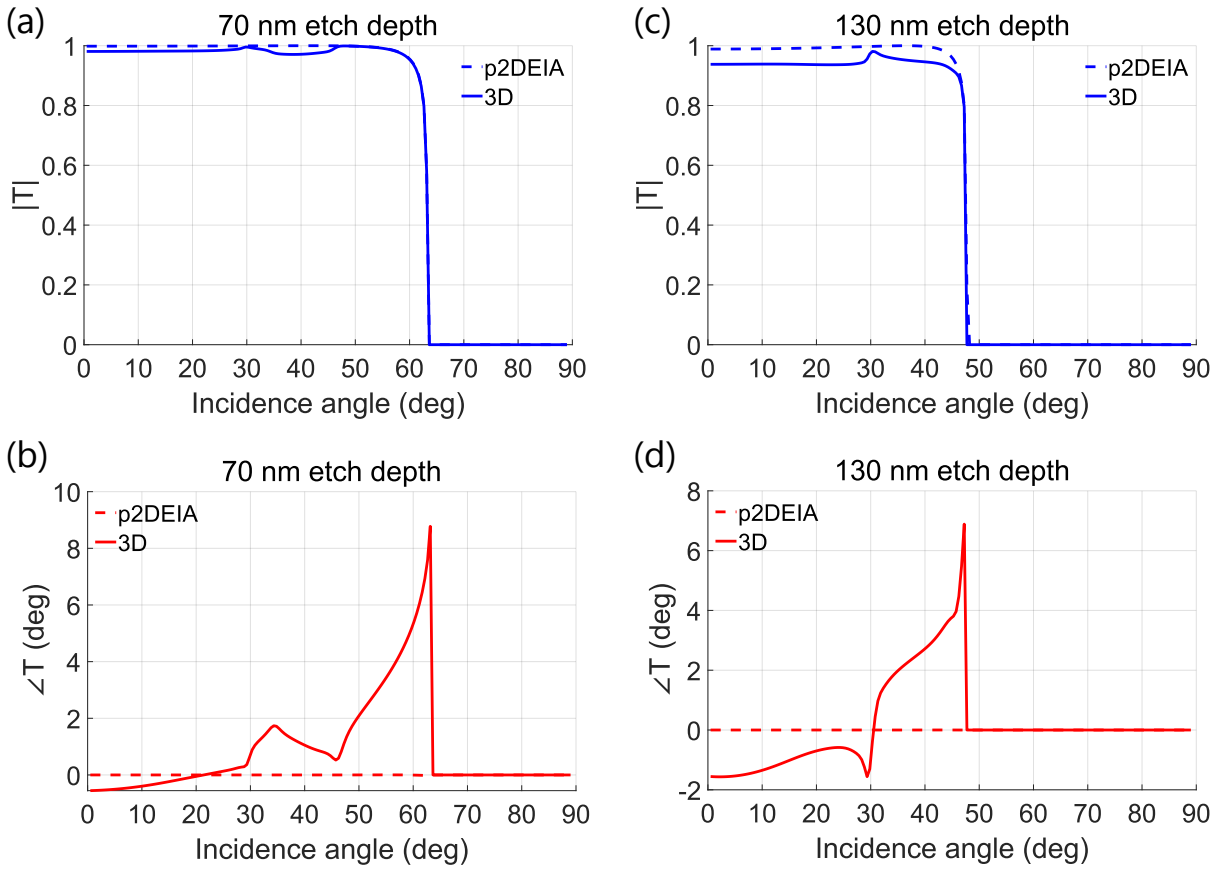

图 S2：**界面透射系数**（a、b）三维 220nm 至 150nm 厚硅界面的透射系数幅度和相位（实线），以及 p2DEIA 模型中相应投影界面的透射系数幅度和相位（虚线）。（c、d）220nm 至 90nm 硅界面的相应图

此外，从图 S2(b) 和 (d) 的相位曲线可以观察到，在临近临界角的区域，由于失配加剧，三维模型与 p2DEIA 之间的相位差异会出现尖峰。然而，在临界角之前，$70\ nm$刻蚀高度的相位误差曲线小于 $130\ nm$情况，这进一步证明了采用 $70\ nm$刻蚀高度能够获得更高的准确性。

为求完整，我们也在图 S3 中展示了反射系数的类似数据。由于这是一个通常为前馈系统的系统，这一点不太重要；不过显然，刻蚀深度较浅时，反射幅度较小。

在实际设计中，波会与输入端和

输出端之间的多个界面发生相互作用。每次相互作用都会造成一定的功率损耗和轻微的相位旋转，累积起来会导致整体损耗因子和相位误差。

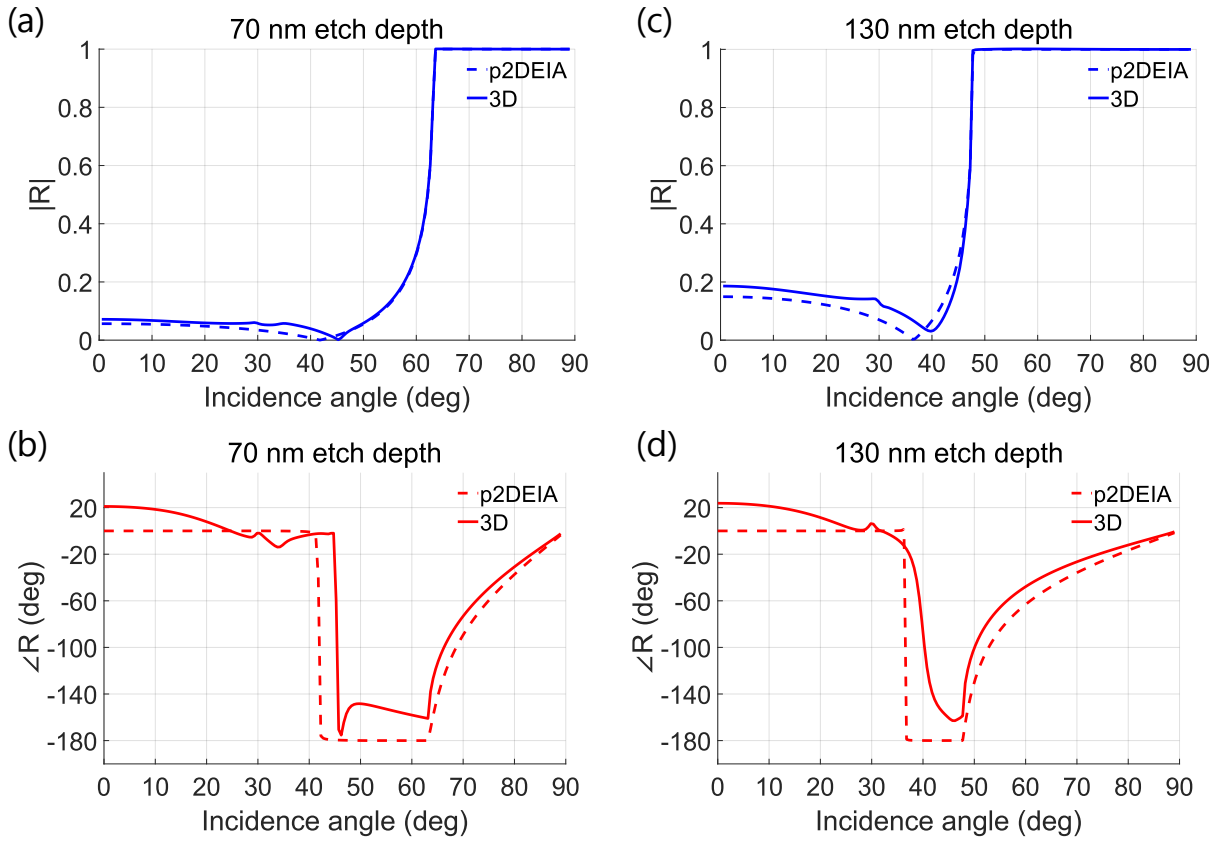

图 S3：**界面反射系数** (a,b) 三维 220nm 至 150nm 厚硅界面的反射系数幅值和相位（实线），以及 p2DEIA 模型中对应投影界面的反射系数幅值和相位（虚线）。(c,d) 220nm 至 90nm 硅界面的对应曲线

散射误差归一化

我们希望为器件考量一种有意义的相对误差。考虑将实际实验得到的 $N \times N$ 矩阵 $\mathbf{T}_{\text{exp}}$ 与相应的目标矩阵 $\mathbf{T}_{\text{targ}}$进行比较时所定义的误差。平方误差的一种自然度量是弗罗贝尼乌斯范数的平方，即 $\|\mathbf{T}_{\text{exp}} - \mathbf{T}_{\text{targ}}\|_F^2$。然而，使用这种度量难以比较不同大小的矩阵，因为对于数值相近的矩阵，较大的矩阵自然会比小矩阵具有更大的范数。另一方面，

这些矩阵的数值很可能并不相近，因为无源散射矩阵的元素值自然会随着矩阵尺寸增大而减小。这可能会导致较大矩阵的误差看起来更小。我们希望找到一种误差度量，能够克服这些问题，对不同器件进行比较。

我们来考虑矩阵元素 $T_{i,\ j}$的平均相对平方误差 。平均相对平方误差为 $\|\mathbf{T}^{\text{exp}} - \mathbf{T}_{\text{targ}}\|_F^2/(N^2\Delta_{\text{typ}}^2)$，其中 Δ_{typ}是某个典型的复数距离值， N^2 是矩阵中的元素个数。在物理学和工程学的许多场合，典型距离值取为目标 值本身。然而，这种做法会过分强调较小目标值的重要性。事实上，如果某个目标元素为零，那么任何实验实现都会具有无限大的相对误差，即使其实际表现相当不错。相反，我们注意到，感兴趣的是酉（或接近酉）的目标矩阵，其元素的幅度约为 $|\mathbf{T}_{i,\ j}| \approx 1/\sqrt{N}$。由于我们关注的是复数 误差，相位相差 180 度的实验实现值，其典型绝对误差距离 Δ_{typ} 将为 $2/\sqrt{N}$。因此，平均相对平方误差变为误差$^2 = \|\mathbf{T}^{\text{exp}} - \mathbf{T}_{\text{targ}}\|_F^2/(4N)$。

—

—

该指标具有以下有用性质。假设目标矩阵$\mathbf{T}^{\text{targ}}$是酉矩阵。如果$\mathbf{T}^{\text{exp}}$ =-$\mathbf{T}^{\text{targ}}$，那么误差$^2 = 1.00$。如果$\mathbf{T}_{\text{exp}}$ = **0**，那么误差$^2 = 0.25$。如果$\mathbf{T}^{\text{exp}}$ = $\mathbf{T}^{\text{targ}}$，那么无论是否存在零项，误差$^2 = 0.00$。

制造

光子芯片采用 AMF 180 nm SOI工艺制造，埋氧层厚度为 2 μm。光子信号通过宽度为 500 nm 、高度为 220 nm 的单模波导在芯片上进行传输。光通过光栅耦合器耦合至芯片，耦合效率约为 40%。AMF 工艺还提供响应度约为 0。8 A/W 的光电二极管，用于监测输出光功率。

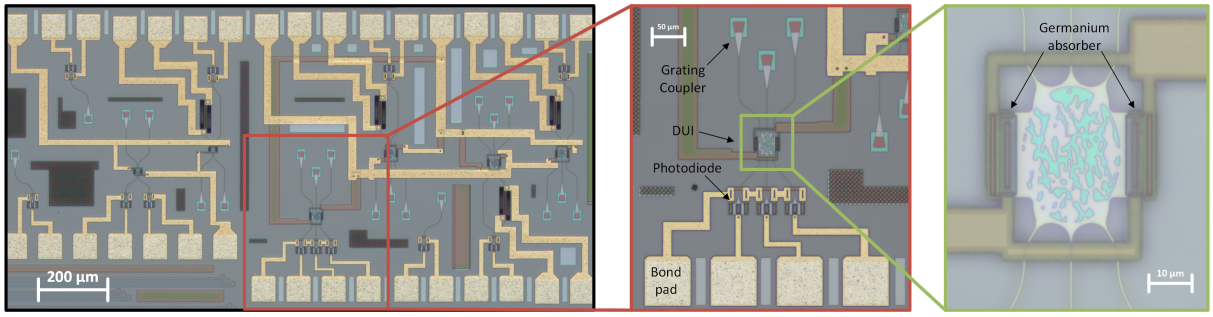

图 S4：物理实验的显微照片，显示了若干个 2×2 和 3×3 卷积核以及各种校准结构。 3×3 卷积核已突出显示。

测量

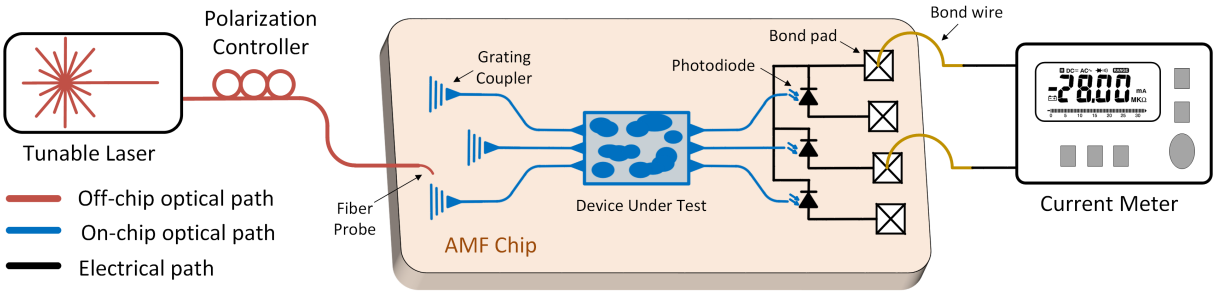

图 S5：测量装置。使用可调谐激光器照射多个光栅耦合器中的一个。光经过逆向设计结构。最后，使用电流计测量三个光电二极管各自的功率。

使用可调谐激光器测试所制造的逆向设计器件的性能。如图 S5 所示，红外光通过光纤探针输入片上光栅耦合器。使用偏振控制器调节光偏振，使其与光栅耦合器的偏振响应相匹配，并最大化耦合至片上波导的光功率。在输出端，采用光电二极管将光功率转换为电流，随后使用电流计测量该电流。

三维与 p2DEIA 结构的计算资源分析

一个自然而然的问题是：与全波三维模型相比，采用 p2DEIA 进行逆向设计方法能在定量上带来哪些收益？由于对大型结构进行逆向设计在计算上可能并不可行，我们将采用启发式分析。逆向设计采用伴随法，通过一系列仿真计算梯度；在这些仿真中，“输入”和“输出”端口均受到照明。无论优化的是 p2DEIA 模型还是三维模型，对于给定的散射参数目标，每次优化迭代所需的仿真次数相同。因此，初步比较执行优化所需的计算量时，只需比较两种情况下求解光场的成本即可。不过，这种比较只是近似的，因为还有许多细节会影响两种情况下的收敛过程。

本节比较求解全波三维模型和 p2DEIA 模型中的光场所需的计算资源。与本文研究中的情况一样，我们考虑一种经过设计、用于执行 10×10 向量-矩阵乘积的硅光子学超结构（见正文）。我们通过考虑输入端口上一个随机的模式输入向量来考察该结构。我们采用有限元法（FEM），使用 Intel(R) Xenon(R) Gold 6130 CPU @ 2.10GHz（16 个芯层，随机存取存储器为 225GB），在设计波长 $\lambda_0 = 1.525\ \mu m$ 处仿真时谐光场。求解线性系统时，我们使用 COMSOL 5.6 中的广义最小残差法（GMRES）迭代求解器（*35*）和并行直接求解器 PARDISO（*36*）。表 1 列出了域离散化、计算资源以及求解线性系统所需时间的相关信息，详见下文：

- **网格体积（面积）：** 三维（p2DEIA）结构的仿真域体积（面积）。

- **网格 ce lls（阶数）**：离散仿真域的单元总数。对于 th e

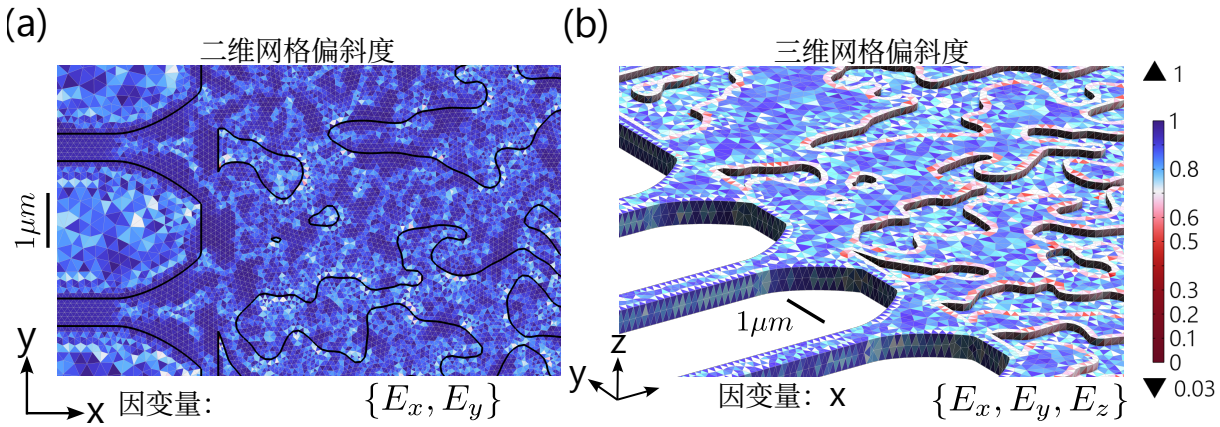

图 S6：**网格偏斜度**（a） 10×10 硅光子学超结构的 p2DEIA 模型网格图；（b）对应的三维模型网格图

三维仿真中，大多数单元为四面体。对于 p2DEIA 仿真，单元为三角形。此处的阶数 指与各个单元相关联的形函数阶数。

• **平均质量**：如图 S6 所示，我们将网格偏斜度作为品质因子。它量化了单元相对于适用于离散化的理想参考单元的变形程度。偏斜度评价值介于 0（差）和 1（好）之间。此处报告仿真域内偏斜度评价值的平均值。

• **自由度 (DOF)**：表征光场的未知数数量。它取决于作为因变量的电场分量数量、形函数以及单元数量。根据研究中的情况，对于 p2DEIA 仿真，我们考虑 $\{E_x, E_y\}$；对于三维模型，我们考虑全部三个分量，$\{E_x, E_y, E_z\}$。

• **物理内存（RAM）**：线性系统占用的物理内存大小。

• **时间**：求解线性系统所需的时间。

每次随机激励的计算资源分析（有限元法）						
EM 模型（求解器）	Mesh 体积（面积）	单元（阶数）	平均质量	DOF	RAM	Time
p2DEIA（PARDISO）	788 μ米^2	181K (立方)	0.85	1.92M	8.01GB	4.9秒
p2DEIA (GMRES)	788μ米^2	181K (立方)	0.85	1.92M	3.4GB	16分钟
3D (GMRES)	3058μ米^3	926K（立方）	0.63	18.42M	186GB	5小时45分钟
3D（PARDISO）	3058μ米^3	926K（立方）	0.63	18.42M	400GB	N.A.

数据显示，p2DEIA 模型能够通过在每次迭代中快速评估光场，同时减少计算资源的使用，显著加快逆向设计过程。