无源空间中的学习：端到端多阶端口哈密顿宏模型构建

Benchao Zhu, Keren Zhu*复旦大学集成电路与微纳电子学院，上海，中国
zhubc25@m.fudan.edu.cn, krzhu@fudan.edu.cn

arXiv:2611.28778v2 [cs.AR] 24 Nov 2026

摘要—**互连线和供电网络的紧凑型无源状态空间宏模型对于稳定的瞬态信号完整性和电源完整性仿真至关重要。然而，传统流程需要针对每个布局进行成本高昂的全波电磁（EM）仿真，而传统拟合和以往的学习代理模型都依赖事后拟合或无源性扰动，这些方法会降低精度且无法保证收敛。在本文中，我们提出了一种端到端宏建模框架，可直接在无源空间中进行学习，将物理布局图映射为多阶端口哈密顿（pH）宏模型。具体而言，我们的网络将每个布局图编码为无约束矩阵块，并通过完整的前缀嵌套式 pH 参数化将其组装起来，从结构上保证连续频率无源性和稳定性，并通过一次前向传播中的直接状态切片生成有效的 k阶状态空间宏模型。为在无需参考实现的情况下训练所有阶数，我们直接基于 S参数对网络进行监督，并采用渐进式阶数增长课程。在来自一个包含 270k 个结构的先进封装数据集的 27k 个未见测试结构上进行评估时，与基于真实电磁响应拟合的成熟宏模型构建工具相比，我们的框架平均将 S参数误差降低**42%（**最高**65%），**在**459,000**个导出模型中实现零无源性违规，并将宏模型生成速度提升**11,000–39,000×**倍（每个布局**36–85**毫秒**）。

I. 引言

紧凑型状态空间宏模型广泛用于近似互连线和供电网络（PDNs）在瞬态信号完整性与电源完整性（SI/PI）仿真中的端口行为 [6]。可靠的时域分析要求这些模型以较少的状态实现高精度 [7]，同时保持严格无源，因为即使轻微违反无源性，也可能在连接无源终端时触发仿真发散 [8]。

传统的宏模型构建遵循图1（上方）所示的三阶段流程。首先，使用 HFSS 等全波电磁（EM）求解器 [9] 计算离散频率处的散射参数（S参数）。接着，对这些频率样本拟合指定阶数的有理模型（例如，采用向量拟合 [10]）。最后，通过无源性强制对拟合模型进行扰动，直到其在给定容差范围内满足无源性判据 [11], [12]。

然而，该流程的效率和可靠性受到两个瓶颈的限制。首先，对每种结构进行全波 EM 求解需要数分钟到数小时，而且每次布局修改后都必须重新求解。其次，无源性并非拟合模型的固有属性，而是通过事后扰动强加的，这会降低精度，并可能导致无法收敛。如图1（右侧）所示，即使经过无源性强制，模型仍可能保持非无源，从而导致瞬态仿真不稳定，即使未驱动端口保持开路也是如此。

以往的工作通常只缓解其中一个瓶颈，而未解决另一个问题。非学习型无源建模技术能够确保各频率下的无源性，但仍受限于电磁仿真成本或参数维度。当已有 RLC网表时，模型降阶可以保持无源性；然而，针对采样响应的无源拟合与细化方法（例如固定极点凸拟合 [13], 端口哈密顿（pH）拟合 [14], 和谱因子细化 [15], ）对于每个新布局仍需要进行一次全波电磁仿真。参数化宏模型 [16], [17] 和参数化端口哈密顿降阶 [18]能够在设计空间内构建无源模型，但依赖于固定基函数或插值方案，而这些方法只能扩展到少数几个参数。

另一方面，基于学习的方法可以绕过对未见布局进行电磁仿真，但无法保证连续频率无源性。基于学习的代理模型根据布局直接预测 S 参数或 Z 参数 [1]–[4]；然而，即使输出层对奇异值施加了界限 [19]，其离散频率输出仍需进行有理拟合和无源性强制。神经传递函数（neuro-TF）模型进一步通过根据设计变量预测有理传递函数来消除后续拟合 [5], ，但即使具有稳定极点并对每个样本的奇异值进行缩放 [20],，无源性至多也只能在采样频率处成立，而无法覆盖整个连续频谱。

要解决这些瓶颈，必须将学习过程本身约束在无源空间内。为此，我们提出一种端到端宏模型框架（图1，下方），将每种布局表示为图，并采用图神经网络（GNN）与多层感知机（MLP）直接组装端口哈密顿实现 [21]，无需测试时电磁仿真、有理拟合或事后强制。此外，为了适用于状态阶数需求各不相同的结构，而不必为每种阶数单独建立模型或进行事后降阶，每次前向传播都会输出具有多个状态数量的嵌套实现。据我们所知，这是首个能够直接根据布局学习嵌套连续频率无源宏模型的方法，且这些布局在工艺、布线拓扑、回流路径以及数十种几何和材料参数方面各不相同。我们的主要贡献如下：

- **完整的无源参数化**。我们将无约束的网络输出以可微方式映射到全部可容许的端口哈密顿实现集合，从而能够直接在无源空间内进行优化，而无需损失惩罚或事后扰动。
- **嵌套式多阶实现**。我们将能量矩阵固定为单位矩阵，并对耗散矩阵进行联合分解，使每个前导k状态切片都保持无源，且不受追加状态的影响，从而在一次前向传播中生成阶数为 $k \in \{k_{\text{最小值}} \ldots, k_{\text{最大值}}\}$ 的宏模型。
- **S参数阶数增长课程学习**。我们直接基于S参数进行训练，无需参考状态空间拟合；同时将状态逐步增长与几何平均多阶损失相结合，避免低阶残差主导高阶细化。

我们在一个包含 27 万个经电磁仿真得到的先进封装结构的数据集上评估了所提出的框架，涵盖信号通道、垂直过渡结构和供电网格。实验结果表明，我们的网络可在 36–85 毫秒内根据未见布局生成无源宏模型（比基于电磁仿真的流程快 11，000–39，000× 倍），在全部 459，000 个导出的

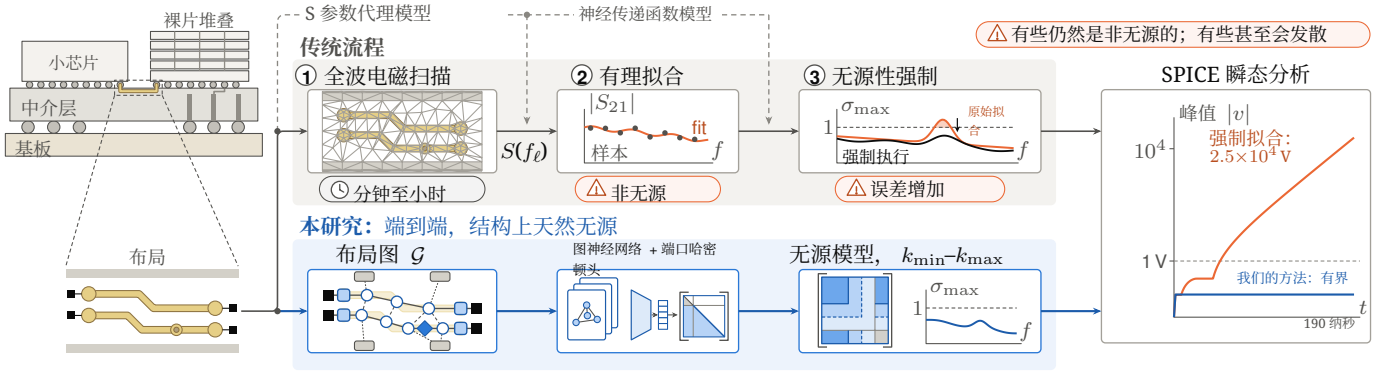


图1. 传统宏模型建模流程与所提出的端到端宏建模框架的比较。上：传统建模依次执行 (1) 全波电磁扫描、(2) 有理拟合，以及 (3) 基于扰动的无源性强制，将最大奇异值限制在 $\sigma_{\max}$。虚线表示以往绕过扫描 [1]–[4]，或同时绕过扫描和拟合 [5] 的学习代理模型，但仍需要进行无源性强制。下：在电磁仿真结构上训练后，我们的网络可通过一次前向计算，将未见过的布局图直接映射为 $k_{\min}$ 至 $k_{\max}$ 个状态的嵌套无源稳定实现。右：瞬态 ngspice 阶跃响应（通过 50 Ω 电阻向端口 1 输入 1 V 阶跃，端口 2–8 开路）。强制拟合的 32 状态模型仍非无源，并在 190 ns 内发散至 2.5×10^4 V，而我们的模型始终保持有界。

测试实现中均未出现无源性违规；与基于真实电磁响应拟合的 MFT-NNLS、MATLAB RF Toolbox 和 Ansys Genequiv 相比，S 参数平均误差平均降低 42%（最高可达 65%）。与去噪多次求解电磁参考进行的评估进一步表明，跨布局学习能够减轻对单次求解数值波动的敏感性，而逐结构拟合器会部分吸收这些波动。

II. 预备知识

本节回顾状态空间宏模型，介绍端口-哈密顿系统，并正式陈述问题。

A. 状态空间宏模型

考虑一个以阻抗形式表示的线性时不变（LTI） m端口结构，其 n状态实现为：

$$\dot{x} = Ax+Bu, \quad y = Cx+Du, \quad Z(s) = C(sI-A)^{-1}B+D, \tag{1}$$

其中， $x \in \mathbb{R}^n$ 为状态向量，输入 $u \in \mathbb{R}^m$ 和输出$y \in \mathbb{R}^m$ 分别汇集端口电流和电压，而 $u^{\mathsf{T}}y$ 表示进入端口的功率。对于参考阻抗$Z_0 > 0$，散射矩阵由凯莱变换给出：

$$S(s) = \big(Z(s) - Z_0 I\big)\big(Z(s) + Z_0 I\big)^{-1}. \tag{2}$$

对于零输入（$u = 0$），若每条状态轨迹在所有 $t \geq 0$ 下均保持有界（即 A 的所有特征值实部均非正，且任何虚轴特征值都是半单的），则实现 (1) 是李雅普诺夫稳定的；若 A 的所有特征值都严格位于开左半平面，则系统是渐近（赫尔维茨）稳定的。从零初始状态（$x(0) = 0$）出发，若净吸收能量满足 $\int_0^t u(\tau)^{\mathsf{T}}y(\tau)\, d\tau \geq 0$，对所有输入 u 和所有 $t \geq 0$ 均成立，则系统是无源的。对于实实现，无源性等价于在频域中 $Z(s)$ 为正实，或 $S(s)$ 为有界实；也就是说，它在 $\mathrm{Re}(s) > 0$ 内解析，且 $\|S(s)\|_2 \leq 1$（这意味着对所有 $\omega \in \mathbb{R}$，$\sigma_{\max}(S(j\omega)) = \|S(j\omega)\|_2 \leq 1$）。最小实现的无源性意味着李雅普诺夫稳定性；若 $Z(s)$ 没有纯虚轴（无损）极点，则还意味着赫尔维茨稳定性。模型阶数定义为状态维数 n；对于最小实现，该维数等于 $Z(s)$ [22] 的麦克米兰次数。

B. 端口哈密顿系统

线性端口哈密顿（pH）系统 [21] 是如下形式的 n状态实现：

$$\dot{x} = (J-R)Qx+(G-P)u, \quad y = (G+P)^{\mathsf{T}}Qx+(D_0+N)u, \tag{3}$$

其参数矩阵满足

$$J = -J^{\mathsf{T}}, \quad N = -N^{\mathsf{T}}, \quad Q = Q^{\mathsf{T}} \succ 0, \quad W = \begin{bmatrix} D_0 & P^{\mathsf{T}} \\ P & R \end{bmatrix} \succeq 0. \tag{4}$$

这里， $Q \in \mathbb{R}^{n\times n}$ 是能量矩阵， $J \in \mathbb{R}^{n\times n}$ 是互联矩阵，$W \in \mathbb{R}^{(m+n)\times(m+n)}$ 是由对称的端口直通项 $D_0 \in \mathbb{R}^{m\times m}$、状态耗散 $R \in \mathbb{R}^{n\times n}$和端口—状态耦合 $P \in \mathbb{R}^{n\times m}$构成的耗散矩阵。反对称矩阵 $N \in \mathbb{R}^{m\times m}$ 表示无损直通项，而 $G \in \mathbb{R}^{n\times m}$ 无约束。对于二次哈密顿量 $\mathcal{H}(x) = \frac{1}{2}x^{\mathsf{T}}Qx$， J 和 N 的反对称性给出功率平衡关系 –

$$u^{\mathsf{T}}y - \dot{\mathcal{H}}(x) = \begin{bmatrix} u \\ Qx \end{bmatrix}^{\mathsf{T}} W \begin{bmatrix} u \\ Qx \end{bmatrix} \geq 0. \tag{5}$$

从 $x(0) = 0$积分可得 $\int_0^t u(\tau)^{\mathsf{T}}y(\tau)\, d\tau \geq \mathcal{H}(x(t)) \geq 0$，从而确保无源性。此外，由于 $Q \succ 0$和 $R \succeq 0$，令 $u = 0$可得 $\dot{\mathcal{H}}(x) = -x^{\mathsf{T}}QRQx \leq 0$。因此，所有满足(3)–(4)的实现，无论是否为最小实现，都是无源且李雅普诺夫稳定的；当 JQ 没有特征向量属于ker(RQ)时（特别是当 $R \succ 0$时），它们还是赫尔维茨稳定的。

反过来，每个具有最小实现（A, B, C, D）的无源系统，都存在一个阶数相同的等价端口哈密顿实现 n [23]。具体而言，可以通过选择 $Q \succ 0$ 作为 Kalman–Yakubovich–Popov（KYP）不等式 [23] 的一个解，保留原有的状态坐标。因此，通过（3）–（4）对宏模型进行参数化，可在构造时保证无源性和稳定性，同时涵盖所有最小无源实现。

C. 问题陈述

设计族中的每种物理结构都表示为布局图 $\mathcal{G}$，用于编码布局几何、布线拓扑、回流路径以及材料和叠层参数等结构与物理属性。我们将多阶宏模型建模问题表述如下。

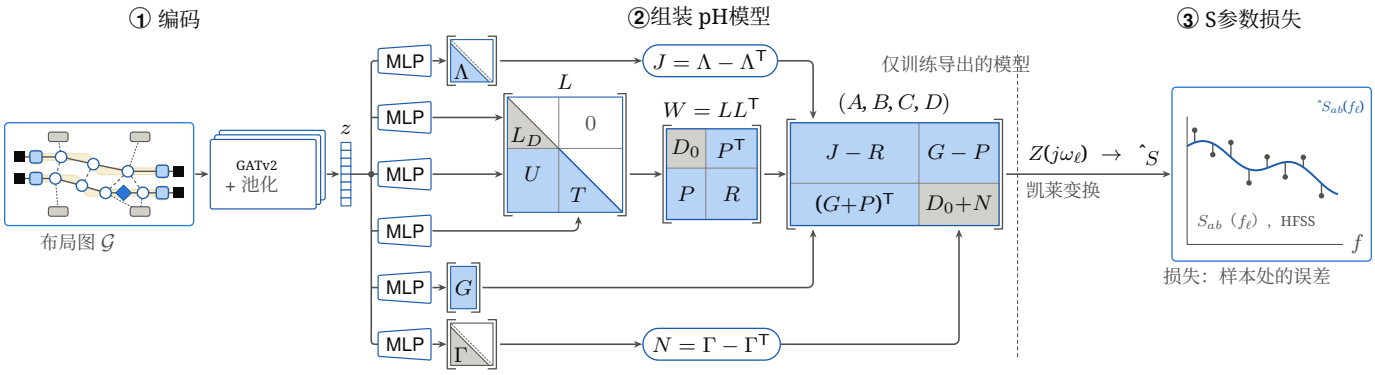


图 2. 所提出的多阶无源宏模型框架概览。(1) GATv2编码器将布局图 $\mathcal{G}$ 映射为潜向量 z。(2) 并行的多层感知机预测无约束块（Λ，L_D，U，T，G，Γ），并以确定性方式组装出一个具有 $k_{\max}$ 个状态的端口哈密顿实现（A，B，C，D），其中包含 $Q = I$（灰色：阶数共享的端口块；蓝色：状态索引块）。(3) 训练期间，对 $Z(j\omega_\ell)$ ($\omega_\ell = 2\pi f_\ell$) 求值并应用凯莱变换，得到 $\hat{S}$，以便直接使用 HFSS 样本进行 S 参数监督。

问题1（多阶无源宏模型化）。给定一个由 m 端口结构组成的族、阶数范围 $\{k_{\min}, \dots, k_{\max}\}$，以及一组由布局图与对应真实频率样本 $\{S(f_\ell)\}_{\ell=1}^{N_f}$ 构成的训练集，学习一种映射，为族中的任意布局图 $\mathcal{G}$ 分配一组阶数为 $k_{\min}$ 至 $k_{\max}$ 的宏模型 $\{\mathcal{M}_k(\mathcal{G})\}_{k=k_{\text{最小}}}^{k_{\max}}$，使得：(i) 每个 $\mathcal{M}_k(\mathcal{G})$ 在所有频率下均为无源且李雅普诺夫稳定；并且 (ii) 其散射响应能够准确逼近该结构的真实 S 参数。

III. 所提出的框架

图 2 展示了所提出的框架。图神经网络编码器将布局图 $\mathcal{G}$ 映射为嵌入向量 z（第 III-D 节），随后并行的多层感知机预测无约束原始块。基于完整的单阶pH参数化（第 III-A 节），我们的嵌套式构造以确定性方式组装出一个具有 $k_{\max}$ 个状态的端口哈密顿实现，其前 k 个状态可在一次前向传播中直接构成适用于所有 $k \in \{k_{\min}, \dots, k_{\max}\}$ 的无源宏模型（第 III-B 节）。训练期间，在采样频率处对组装后的模型求值，并通过阶数增长课程直接使用真实 S 参数进行监督（第 III-C 节）。

A. 完整单阶*pH*参数化

神经网络自然会生成无约束的实值输出，而有效的 n阶pH实现（3）必须满足（4）中的对称性和定性条件。为了在固定阶数 n 的情况下弥合这一差距，同时不牺牲表示能力，我们构造了一个从无约束原始块到完整可容许pH矩阵集合的可微映射。

条件（4）对 J、N和 Q施加了各自的约束，同时还对 W 施加了耦合的半正定条件（其对角块 D_0 和 R 本身也必须是半正定的）。借助标准的反对称分解和三角分解，我们预测严格下三角块 $\Lambda \in \mathbb{R}^{n\times n}$和 $\Gamma \in \mathbb{R}^{m\times m}$、下三角块 $L_D \in \mathbb{R}^{m\times m}$ 和$L_R \in \mathbb{R}^{n\times n}$，以及一个下三角块 $L_Q \in \mathbb{R}^{n\times n}$，其正对角元素经过Softplus激活，从而形成

$$J = \Lambda - \Lambda^\mathsf{T}, \quad N = \Gamma - \Gamma^\mathsf{T}, \quad Q = L_Q L_Q^\mathsf{T}, \\ D_0 = L_D L_D^\mathsf{T}, \quad R = L_R L_R^\mathsf{T}. \tag{6}$$

通过反对称分解和 Cholesky 型分解，(6) 覆盖所有可行的 (J、N、Q、D_0、R)。

要构造完整的耗散矩阵 W，不能省略非对角的端口—状态耦合项 P。虽然设置 $P = 0$ 可直接确保 $W = \mathrm{diag}(D_0, R) \succeq 0$，但这会迫使 $Z(s) - D$ 在所有频率下都为正实，满足 $\frac{1}{2}(Z(j\omega)+Z(j\omega)^\mathsf{H}) \succeq D_0$。这排除了常见的无源结构，例如电阻随频率增加的 RL 单端口 $Z(s) = R_1 + R_2\tau s/(1+\tau s)$ (R_1, R_2, $\tau > 0$)。

为保留所有有效耦合，我们采用半正定分块矩阵的压缩刻画 [24]：给定 $D_0 = L_D L_D^\mathsf{T}$ 和 $R = L_R L_R^\mathsf{T}$，$W = \mathrm{diag}(L_D, L_R)\begin{bmatrix} I & K^\mathsf{T} \\ K & I \end{bmatrix}\mathrm{diag}(L_D, L_R)^\mathsf{T} \succeq 0$ 当且仅当 $P = L_R K L_D^\mathsf{T}$ 对于某个压缩算子 $\|K\|_2 \le 1$成立。因此，我们预测一个无约束的 $\tilde{K} \in \mathbb{R}^{n\times m}$，并将其归一化为

$$P = L_R K L_D^\mathsf{T}, \qquad K = \tilde{K} / \max\big(1, \|\tilde{K}\|_2\big), \tag{7}$$

这在保证对任意网络权重均有 $W \succeq 0$ 的同时保持压缩算子不变。

结合一个无约束的 $G \in \mathbb{R}^{n\times m}$，(3) 给出了 n 状态实现

$$A = (J - R)Q, \qquad B = G - P, \\ C = (G + P)^\mathsf{T} Q, \quad D = D_0 + N. \tag{8}$$

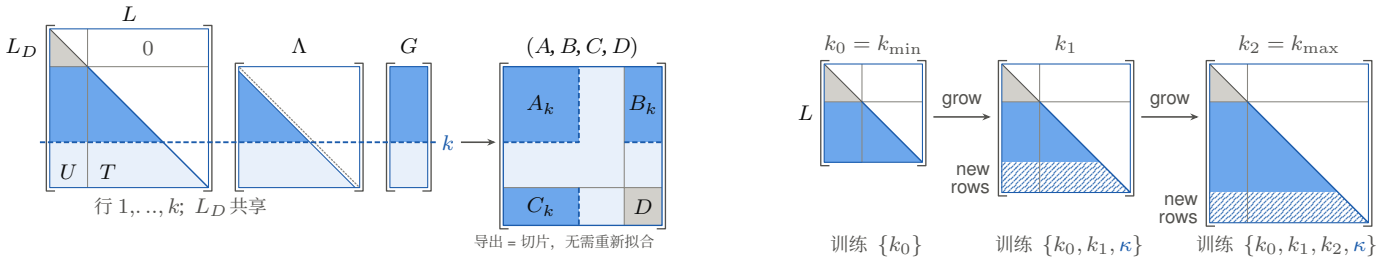
命题 1 (无源性与完备性)。*(a)* 对于任意无约束的原始块，(6)–(8) 满足 *pH* 条件 (3)–(4)，从而保证在所有频率下均具有无源性和李雅普诺夫稳定性。*(b)* 反之，每个 n 状态端口哈密顿实现 (3)–(4) 都具有(6)–(8)形式的参数化。

证明：(a) 根据 (6)，J 和 N 是反对称的，并且由于 L_Q 的对角元素为正，$Q \succ 0$。由 (7) 中的 $\|K\|_2 \le 1$ 可知，W 的分块因式分解保证了 $W \succeq 0$。因此，(8) 满足 (3)–(4)，这意味着系统具有无源性和李雅普诺夫稳定性（第 II-B 节）。(b) 给定任意端口哈密顿实现 (3)–(4)，令 G 保持不变，令 Λ和 Γ分别为 J 和 N的严格下三角部分，令 L_Q 为 Q 的 Cholesky 因子（对其对角线应用 Softplus 的逆变换），并令 L_D，L_R为 D_0，R的下三角因子。由于 $W \succeq 0$，存在满足 $P = L_R K L_D^\mathsf{T}$ [24]的压缩映射 K。因此，(6)–(8) 可恢复 (3)。 ∎

由于命题 1 对任意多层感知机输出均成立，网络可以直接在完全无源空间中学习，无需损失惩罚或事后强制施加无源性。

B. 嵌套多阶 *pH* 参数化

虽然第 III-A 节保证了固定阶数 n 下的无源性，但问题 1 要求宏模型涵盖多个阶数 $k \in \{k_{\min}, \dots, k_{\max}\}$。我们不为每个阶数分别训练独立网络，而是寻求一个单一的 $k_{\max}$ 状态实现（A，B，C，D），从中可通过切片提取每个 k 阶模型，即取其前 k 个状态，得到前导



(a) 嵌套 pH 参数化　　(b) 增长课程

图 3. 多阶状态切片与渐进式训练。(a) 嵌套 pH 参数化：切取状态索引块的前 k 行（Λ、U、T、G），同时保留共享端口块（L_D、Γ），即可得到有效的 k 状态无源实现（A_k、B_k、C_k、D）。(b) 阶数增长课程学习（以 $\bar{g} = 2$ 为例）：在阶段 $g = 0$、……、$\bar{g}$（$k_0 = k_{\min} < \cdots < k_{\bar{g}} = k_{\max}$）中，逐步追加新初始化的状态行，同时不改变低阶模型，并联合优化锚定阶数和随机采样的中间阶次 κ。

子实现（A_k，B_k，C_k，D），其中 $A_k \in \mathbb{R}^{k\times k}$、$B_k \in \mathbb{R}^{k\times m}$ 和 $C_k \in \mathbb{R}^{m\times k}$。具体而言，我们要求满足两个性质：（i）前缀无源性，确保每个（A_k，B_k，C_k，D）在一次前向传播中均为无源且稳定；（ii）前缀不变性，在追加超过 k 的状态时保持（A_k，B_k，C_k，D）不变，从而能够在训练过程中渐进式地增长状态数量（第 III-C 节）。

然而，$n = k_{\max}$ 处的单阶参数化 (6)–(8) 不满足这两项性质。首先，对于一般的 $Q \neq I$，$A = (J - R)Q$ 的主子矩阵不保证保持稳定（例如，对于 $J = \left[\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right]$、$R = 0.1I$，以及 $Q = \left[\begin{smallmatrix} 1 & 0.9 \\ 0.9 & 1 \end{smallmatrix}\right]$，$A$ 为赫尔维茨稳定，其特征值为 $-0.1 \pm 0.43j$，但 $A_{11} = +0.8 > 0$），这违反了前缀无源性。其次，(7) 中的归一化对所有 $k_{\max}$ 行取谱范数 $\|\tilde{K}\|_2$。只要 $\|\tilde{K}\|_2 > 1$，追加状态就会重新缩放 P 的前 k 行，从而破坏前缀不变性。

为解决这两个问题，我们将能量矩阵固定为 $Q = I$，并以联合耗散矩阵 W的单一下三角分解取代单独的 $D, 0, R$ 因子和谱归一化，并将 m-端口块排在 $k_{\max}$-状态块之前。具体而言，多层感知机预测严格下三角的 $\Lambda \in \mathbb{R}^{k_{\max}\times k_{\text{最大值}}}$和$\Gamma \in \mathbb{R}^{m\times m}$、下三角的 $L_D \in \mathbb{R}^{m\times m}$ 和 $T \in \mathbb{R}^{k_{\max}\times k_{\text{最大值}}}$，以及无约束的 U，$G \in \mathbb{R}^{k_{\max}\times m}$。利用 $J = \Lambda - \Lambda^{\mathsf{T}}$ 和$N = \Gamma - \Gamma^{\mathsf{T}}$，我们构造

$$W = \begin{bmatrix} D_0 & P^{\mathsf{T}} \\ P & R \end{bmatrix} = LL^{\mathsf{T}}, \qquad L = \begin{bmatrix} L_D & 0 \\ U & T \end{bmatrix}, \tag{9}$$

$$A = J - R, \quad B = G - P, \quad C = (G + P)^{\mathsf{T}}, \quad D = D_0 + N, \tag{10}$$

其展开式为 $D_0 = L_D L_D^{\mathsf{T}}$、$P = UL_D^{\mathsf{T}}$和 $R = UU^{\mathsf{T}} + TT^{\mathsf{T}}$。这一构造无需谱归一化，即可对任意原始块保证 $W \succeq 0$，同时保留所有可表示的无源传递矩阵，正如命题 2(b)所确立的那样。

由于 L是下三角矩阵，且端口块排在首位，因此 W 的前导块仅由 L 的前导行决定。具体而言，L的前 $m+k$行在第 $m+k$列之后均为零，因此

$$(LL^{\mathsf{T}})_{[1:m+k,\,1:m+k]} = L_k L_k^{\mathsf{T}}, \qquad L_k = \begin{bmatrix} L_D & 0 \\ U_k & T_k \end{bmatrix},$$

其中，$U_k \in \mathbb{R}^{k\times m}$ 表示 U 的前 k 行，$T_k \in \mathbb{R}^{k\times k}$表示 T的顺序主子矩阵（$G_k \in \mathbb{R}^{k\times m}$ 和$\Lambda_k \in \mathbb{R}^{k\times k}$ 的定义类似）。由于 Λ和 T为下三角矩阵，(A_k，B_k，C_k，D)仅取决于 Λ、U、T和G的前 k行，同时在所有阶次下共用直通项 D（图 3(a)）。

命题 2（前缀嵌套与完备性）。*(a)* 对于任意无约束原始块以及每个 $k \in \{1, \ldots, k_{\max}\}$，其前导部分子实现 (A_k, B_k, C_k, D)（对应(9)–(10)）在所有频率下均为无源且李雅普诺夫稳定（前缀无源性），并且仅取决于(Λ_k U_k T_k G_k L_D Γ $k+,,,,,)$，对行,,保持 $1 \ldots k_{\max}$ $k \in \{1, \ldots, k_{\max}\}$不变（前缀不变性）。*(b)* 对于每个，任意 k 状态端口哈密顿实现(3)–(4) 都具有与某个 k 状态子实现(A_k, B_k, C_k, D) 相同的传递矩阵 $Z(s)$，该子实现具有(9)–(10)的形式。

证明：(a) 根据式 (9)–(10) 以及上述块三角结构，(A_k, B_k, C_k, D) 满足 $A_k = J_k - R_k$, $B_k = G_k - P_k$,$C_k = (G_k + P_k)^{\mathsf{T}}$ 和 $D = D_0 + N$，其中$J_k = \Lambda_k - \Lambda_k^{\mathsf{T}}$和 $\left[\begin{smallmatrix} D_0 & P^{\mathsf{T}} \\ P_k & R_k^{k} \end{smallmatrix}\right] = L_k L_k^{\mathsf{T}} \succeq 0$为反对称矩阵。因此，$(A_k, B_k, C_k, D)$ 满足式 (3)–(4)，其中 $Q = I_k$，且不依赖于第$k+1$行至第$k_{\max}$行。(b) 给定任意一个具有 $Q = L_Q L_Q^{\mathsf{T}} \succ 0$的 k状态端口哈密顿实现 (3)–(4)，应用状态相似变换 $M = L_Q^{\mathsf{T}}$可保持 $Z(s)$，并得到 $(\tilde{A}, \tilde{B}, \tilde{C}, D) = (L_Q^{\mathsf{T}} A L_Q^{-\mathsf{T}}, L_Q^{\mathsf{T}} B, C L_Q^{-\mathsf{T}}, D)$。利用 $QL_Q^{-\mathsf{T}} = L_Q$，可将其化简为 $\tilde{A} = \tilde{J} - \tilde{R}$、$\tilde{B} = \tilde{G} - \tilde{P}$和 $\tilde{C} = (\tilde{G} + \tilde{P})^{\mathsf{T}}$，其中$\tilde{J} = L_Q^{\mathsf{T}} J L_Q = -\tilde{J}^{\mathsf{T}}$、$\tilde{G} = L_Q^{\mathsf{T}} G$、$\tilde{P} = L_Q^{\mathsf{T}} P$和 $\tilde{R} = L_Q^{\mathsf{T}} R L_Q$。此外，$\tilde{W} = \left[\begin{smallmatrix} D_0 & \tilde{P}^{\mathsf{T}} \\ \tilde{P} & \tilde{R} \end{smallmatrix}\right] = \operatorname{diag}(I_m, L_Q^{\mathsf{T}}) W \operatorname{diag}(I_m, L_Q) \succeq 0$，因此它存在一个下三角分解 $\tilde{W} = L_k L_k^{\mathsf{T}}$，其中 $L_k = \left[\begin{smallmatrix} L_D & 0 \\ U_k & T_k \end{smallmatrix}\right]$。令 Λ_k和 Γ 分别为 J和 N的严格下三角˜　　˜部分，并设 $G_k = G$（第> k行的元素可任意取值），即可恢复 $(\tilde{A}, \tilde{B}, \tilde{C}, D)$，进而通过式 (9)–(10) 得到 $Z(s)$。 ■

根据命题　　2,　　$k \in$，通过直接 $\{k_{\min}, \ldots, k_{\max}\}$ 矩阵切片，一次前向传播即可导出所有阶数的无源宏模型，无需事后进行模型降阶 [7]。

C. S参数上的阶数增长课程学习

虽然式（9）–（10）在结构上保证了无源性和前缀嵌套，但训练必须确保每个切片子实现（A_k、B_k、C_k、D）（$k_{\min} \le k \le k_{\max}$）都能准确捕捉物理端口行为。由于状态空间实现存在相似变换意义下的非唯一性，若直接监督（A、B、C、D），就需要预先拟合并规范化参考实现。相反，我们直接基于可观测 S 参数监督网络，无需对任何训练结构进行有理拟合。具体而言，对于训练结构 i和阶数 k，我们在采样频率 $\{f_\ell\}_{\ell=1}^{N_f}$ 处计算阻抗 $Z_k(j2\pi f_\ell) = C_k(j2\pi f_\ell I_k - A_k)^{-1} B_k + D$，并应用凯莱变换（2）ˆ以获得预测散射矩阵 $S_k(f_\ell)$（为简洁起见，矩阵中省略结构索引 i）。将每个阶数的结果直接与真实电磁响应 $S(f_\ell)$进行比较，我们将每个结构的误差定义为

$$e_{i,k} = \left(\frac{1}{N_f m^2} \sum_{\ell=1}^{N_f} \sum_{a,b=1}^{m} w_\ell \, \rho_\beta\big([\hat{S}_k(f_\ell) - S(f_\ell)]_{ab}\big) \right)^{1/2}, \tag{11}$$

其中 a、b 表示端口索引，$\rho_\beta(z) = h_\beta(\mathrm{Re}\, z) + h_\beta(\mathrm{Im}\, z)$表示应用 Huber函数 h_β ($\beta = 0.01$)，而 $w_\ell \propto 1 + (f_\ell / f_{\max})^2$（归一化为均值为 1，且 $f_{\max} =$ 最大值$_\ell\, f_\ell$）则对高频段进行适度加权，以减轻状态空间梯度的高频衰减。

联合优化多个阶数需要在活动阶数集合 $\mathcal{K} \subseteq \{k_{\min}, \ldots, k_{\max}\}$上汇总其小批量平均误差 $\bar{e}_k$，然而可达到的近似误差会随状态数量显著变化。由于算术平均会使较大的低阶残差主导梯度，并阻碍高阶细化，我们改用加权几何平均:

$$\mathcal{L} = \exp\Big(\sum_{k \in \mathcal{K}} \eta_k \log \bar{e}_k \Big), \qquad \sum_{k \in \mathcal{K}} \eta_k = 1, \tag{12}$$

权重为 $\eta_k > 0$。由于其梯度满足 $\nabla \mathcal{L} = \mathcal{L} \sum_{k \in \mathcal{K}} \eta_k (\nabla \bar{e}_k / \bar{e}_k)$，更新方向会按 η_k对各阶数的相对灵敏度 $\nabla \bar{e}_k / \bar{e}_k$ 加权，从而避免粗略的低阶误差掩盖高阶精度。

利用命题2(a)的前缀不变性，我们通过一个包含 $\bar{g}$ 个阶段、针对状态数量 $k_0 = k_{\min} < k_1 < \cdots < k_{\bar{g}} = k_{\max}$ 的阶数增长课程学习来优化 (12)（图 3(b)），使高阶实现能够建立在已收敛的低阶模型之上。具体而言，阶段 $g = 0$，1，.....，$\bar{g}$ 使用 k_g 个状态训练网络。阶段 0 仅训练 $k_{\min}$状态模型；之后每个阶段都会向 Λ、U、T 和 G 追加 $k_g - k_{g-1}$ 个状态行$g \geq 1$（在 Λ中初始化为零，在 U中初始化为近零，G，并在 T中初始化为对角形式）。为避免每一步都评估全部 $k_g - k_{\min} + 1$ 个阶数，阶段 $g \geq 1$ 设置 $\mathcal{K} = \mathcal{A}_g \cup \{\kappa\}$，将固定的锚定阶数 $\mathcal{A}_g = \{k_0, \ldots, k_g\}$ 与每一步从 $\{k_{\min} + 1, \ldots, k_g - 1\}$ 中均匀采样的阶数 κ 配对。

D. 布局图与编码器

为了预测不同工艺、布线拓扑和基元数量的结构中的原始块（Λ、Γ、L_D、U、T、G），我们将每个结构表示为异构布局图 $\mathcal{G}$，并附加全局叠层参数。$\mathcal{G}$ 中的节点表示端口和布局基元（走线段、过孔、焊盘、参考平面片、电源/接地块和去耦元件），每个节点包含 15 个数值特征（例如位置、尺寸以及过孔/焊盘直径）和 7 个类别特征（例如基元类型、角色、层和网络组）。边用于编码串联连接、层间转换、同层耦合、跨层邻近关系以及端口或参考连接；每条边都包含一个边类型和 8 个数值特征（例如空间偏移、距离、重叠长度、最小间隙和参考标志）。在结构层面，70 个数值参数（例如布局尺寸、层厚度、ε_r和 $\tan\delta$）以及 12 个类别标识符（例如叠层和材料角）用于描述横截面和材料。嵌入后的这些全局特征会与每个节点输入、边输入以及池化图向量拼接。

如图 2 所示，图神经网络编码器将 $\mathcal{G}$ 映射为图嵌入 $z \in \mathbb{R}^{512}$，六个并行的多层感知机将 z 映射为原始块。具体而言，编码器堆叠了三个 GATv2 层 [25]（分别具有 2、2 和 1 个注意力头，每个注意力头包含 256 个通道），并在前两层之后加入 ELU 激活函数和层归一化。节点特征通过均值池化和最大池化进行聚合，然后输入一个采用双曲正切激活函数的两层多层感知机，得到有界嵌入 $z \in \mathbb{R}^{512}$，从而避免初始矩阵块条件不良。基于 z，六个具有 1024 个隐藏单元并采用 SiLU 激活函数的两层多层感知机，输出式 (9)–(10) 中 Λ、Γ、L_D、U、T 和 G 的自由元素。

IV. 实验

A. 实验设置

a) 数据集：我们的数据集包含 27 万个用于模块化封装建模的八端口模块，涵盖三种物理类别，每种类别均按 80/10/10 划分为训练集、验证集和测试集。(1)裸片间信号通道（12 万个结构，包含 85–142 个图节点）由四个子类别组成：硅中介层、硅桥，以及两种重布线层变体 RDL-A 和 RDL-B（通道跨越桥接的参考平面槽）。(2)垂直跃迁（9 万个结构，包含 84–111 个图节点）包括硅通孔（TSV）、微凸点和混合键合。(3) 供电（6 万个结构，包含 65 个图节点）由硅和重布线层中介层上的电源网格单元构成。

真实 S 参数采用 Ansys HFSS 3D Layout 2024 R1 [9] 生成，涵盖 $N_f = 200$ 个频率：信号通道和垂直跃迁为 0.25–50 吉赫兹（$Z_0 = 50\,\Omega$），供电为 0.1–20 吉赫兹（$Z_0 = 1\,\Omega$）。

b) 实现与训练：我们使用嵌套参数化（9）–（10），为每个族在其所有子类别上训练一个统一网络。每个网络在 $\bar{g} = 2$ 增长阶段进行训练，并采用均匀的阶数权重 η_k，将信号通道和供电的锚定阶数（k_0、k_1、k_2）设置为（16、24、32），将垂直跃迁的锚定阶数设置为（8、16、24）。为在保证无源性的同时确保严格的赫尔维茨稳定性，我们将 $R \leftarrow R + \varepsilon I$ 与 $\varepsilon = 10^{-3}$进行移位。具体而言，我们在阶段0中对基础 k_0阶状态模型训练400–600个周期，并在之后两个各持续200–400个周期的阶段中扩展状态空间。模型在 NVIDIA RTX 6000 Ada GPU 上使用 SOAP [26]进行优化，权重衰减为 0.03，指数移动平均衰减为 0.9995；在阶段0中将初始学习率设置为 10^{-3}，在增长阶段设置为 1.5×10^{-4}（Λ多层感知机使用更高的5×），并在验证指标平台期时减半至 10^{-5}。

c) 基线与指标：我们将模型与三种广泛采用的宏模型构建工具进行比较：MFT-NNLS [27], MATLAB RF Tool- box [28], 和 Ansys Genequiv [29], ，这些工具均针对真实电磁响应进行拟合，而我们的网络则直接根据未见布局进行预测。由于 p 跨 $m = 8$个端口共享的极点会产生 mp 个状态（直接进行 k状态拟合时仅有k/m 个极点），对于每个目标阶数 k，我们遍历$p = \lceil k/m \rceil$、$\lceil k/m \rceil + 1$、.....，应用内置的无源性强制和有界实平衡截断 [30], 将模型缩减至 k 个状态（若残余非无源性导致无法使用，则采用标准平衡截断 [31] ），当平均测试误差的改善幅度低于 1% 时停止，并报告最佳 p。我们不纳入以往的 S 参数代理模型 [1]–[4] 和神经传递函数模型 [5]–[20]: 对精确电磁响应进行拟合，已经为代理模型加拟合流程提供了精度上限；而神经传递函数模型仅适用于固定拓扑设计参数。此外，如果不进行事后拟合或强制处理，这两类方法都无法保证连续频率无源性（0.1–7.7 秒）。

我们通过未加权均方根误差 RMS $|\Delta S| = \Big(\frac{1}{N_f m^2} \sum_{\ell,\, a,\, b} | \hat{S}_{ab}(f_\ell) - S_{ab}(f_\ell) |^2 \Big)^{1/2}$ () 评估频域精度，并针对真实电磁标签验证全频谱无源性；如果模型不稳定或满足 $\sup_{\omega \in [0,\infty]} \sigma_{\max}(S(j\omega)) > 1 + \tau$，则在容差 $\tau \in \{10^{-6}, 10^{-4}\}$ 下将其标记为非无源。

B. 主要结果

a) 精度与无源性：表 I 比较了全部九个子族中的宏模型精度与无源性。由于各子族的电气复杂度差异很大（从短混合键合到带槽重布线层通道），因此每个子族均在最小锚定阶数 k 下进行评估；在该阶数下，最精确的基线模型可达到验证集均值 RMS $|\Delta S| \leq 2 \times 10^{-3}$，上限为 $k_{\max}$。

我们的网络相较于每一行的最强基线，平均将均方根 $|\Delta S|$ 降低了42%（相对于每个单独工具的整体平均值降低了49%–76%），在九个子系列中的八个子系列实现了最低平均误差，并在七个子系列中实现了最低最大误差；尽管如此，我们仍能直接从未见过的布局图进行预测，而无需测试时电磁仿真。具体而言，我们的切片模型将平均误差降低了49%–65%】【。

表 I 九个子族之间的宏模型精度与无源性违规对比（每个子族 3,000 个测试结构）。**均方根 $|\Delta S|$**：平均值 / 最大值误差（$\times 10^{-3}$）；#违规：在 $\tau = 10^{-6}$ / 10^{-4}处的非无源计数；**我们的方法 / 最佳方法**：每行相对于最佳基线的平均误差比；底行：平均误差/比值和违规总数。

族	子族	k	MFT-NNLS		MATLAB RF Toolbox		Ansys Genequiv		Ours		我们的方法 / Best
			均方根 $\|\Delta S\|$	#违规数	均方根 $\|\Delta S\|$	#违规次数	均方根 $\|\Delta S\|$	#违规次数	均方根 $\|\Delta S\|$	#违规次数	
信号车道	硅中介层	24	1.87 / 10.6	0 / 0	5.42 / 47.8	126 / 107	2.07 / 15.9	6 / 6	**0.653 / 2.83**	**0 / 0**	0.35
	硅桥	24	0.964 / 4.30	0 / 0	3.06 / 59.1	52 / 33	1.27 / 4.88	0 / 0	**0.489 / 2.03**	**0 / 0**	0.51
	RDL-A	32	3.20 / 17.2	37 / 33	6.49 / 94.2	47 / 42	2.88 / 11.8	3 / 2	**1.20 / 3.53**	**0 / 0**	0.42
	RDL-B	32	2.92 / 19.9	4 / 4	6.86 / 121	350 / 308	2.91 / 12.1	5 / 0	**1.33 / 4.35**	**0 / 0**	0.46
垂直跃迁	TSV	24	1.19 / **7.50**	0 / 0	1.56 / 57.2	41 / 35	**0.925** / 12.7	0 / 0	1.04 / 10.7	**0 / 0**	1.12
	微凸点	16	0.634 / **0.899**	0 / 0	1.33 / 5.94	42 / 1	1.46 / 8.50	10 / 10	**0.363** /1.86	**0 / 0**	0.57
	混合键	8	0.891 / 1.54	0 / 0	1.97 / 19.8	116 / 116	0.970 / 1.71	0 / 0	**0.458 / 1.10**	**0 / 0**	0.51
供电交付	硅中介层网格	24	1.30 / 3.78	0 / 0	2.67 / 7.08	0 / 0	1.74 / 10.4	7 / 5	**1.01 / 3.22**	**0 / 0**	0.78
	RDL中介层网格	24	1.79 / 5.52	43 / 32	2.08 / 52.3	6 / 4	2.33 / 9.07	25 / 22	**0.935 / 1.78**	**0 / 0**	0.52
	平均值 / 总计		1.64 / 7.92	84 / 69	3.49 / 51.6	780 / 646	1.84 / 9.67	56 / 45	**0.831 / 3.48**	**0 / 0**	**0.58**

信号通道的平均误差降低了 49%–65%，微凸块和混合键合的平均误差降低了 43%–49%，供电网格的平均误差降低了 22%–48%；对于 TSV，我们的模型优于 MFT-NNLS 和 RF Toolbox，且与 Ansys 的实现相差不超过 12%，同时最大误差更低（10.7 对比 12.7）。基线方法针对每种结构拟合极点，随后进行事后扰动和截断；而我们的网络则直接在无源空间内进行优化。

在无源性方面，我们的网络在所有子类别中均实现零次违规（0/0）。此外，在 27k 个测试结构的全部 17 个阶数（$k = k_{\min}, \ldots, k_{\max}$）中导出的全部 459,000 个实现均保持稳定，并在 $\omega \in [0, \infty]$ 范围内达到机器精度，满足 $\sigma_{\max}(S(j\omega)) \leq 1$。相比之下，即使采用内置的无源性强制措施，基线方法总计仍留下 56–780 个非无源模型（$\tau = 10^{-6}$，表 I）。这些违规通常在采样频率之间或其上方达到峰值，而离散网格检查会漏掉这些违规；即使是很小的违规，也可能在无源终端下触发不稳定性。

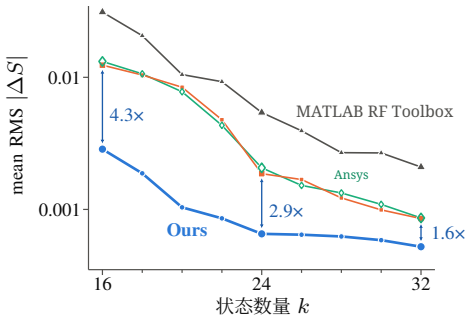


图 4. 硅中介层走线上的多阶测试准确度。箭头表示与最佳基线的平均误差差距及比值。

b) 多阶切片：图 4 展示了从 $k = 16$ 到 32 阶的硅中介层走线测试误差。基线方法需要针对每个目标阶数进行过参数化拟合、无源性强制和平衡截断，而我们的网络只需一次前向传播，即可生成切片后的实现，并在所有阶数下平滑降低误差，表现优于全部三种工具；其中，$k = 24$（6.53×10^{-4}）已优于所有 32 状态基线宏模型。尽管在复杂重布线层走线上，$k = 32$ 仍能将误差降低 54%–58%（表 I），但此处曲线在 $k = 24$ 以上趋于平缓，因为残差已接近单次求解电磁仿真噪声下限，此时继续降低原始标签误差，越来越多地反映了拟合求解器抖动（第 IV-C 节）。

c) 运行时间：表 II 比较了在单个 Xeon Gold 6430 CPU 核心上生成每个结构的宏模型所需时间。而传统流程需要为每个结构耗费 4.9–77 核分钟进行全波电磁仿真（半数 RDL 车道的仿真时间超过一小时），另需 0.1– 7.7 秒进行有理拟合和无源性强制；相比之下，我们的框架可在中位数 36–85 毫秒内，将原始版图映射为其全阶（A, B, C, D）实现（以及所有低阶切片），实现 11,000– 39,000× 倍加速比，并且在针对每种技术类别完成训练后，可将迭代式封装布局探索的周转时间缩短至 100 毫秒以内。

表 II 单个 Xeon Gold 6430 CPU 核心上的每结构宏模型生成时间。EM Solve 报告各子类别的均值（范围）；我们的方法报告布局到实现时间的中位数。

族	EM Solve（核心分钟）	我们的方法（毫秒）	加速比
信号通道	54.9（32–77）	85	39,000×
垂直跃迁	12.5（4.9–21）	69	11,000×
供电	18.6（15–22）	36	31,000×

C. 对电磁求解器噪声的鲁棒性

除了提升标准电磁标签上的精度外，基于确定性几何的跨布局无源学习还能增强对全波求解器数值波动的鲁棒性。具体而言，我们从信号通道测试结构中随机抽取一个子集，在扰动网格划分和扫描设置的条件下，对每个结构重复进行 HFSS 求解并取平均，以形成去噪后的 S 参数参考值。针对这一独立参考值重新评估现有宏模型（不进行重新训练或重新拟合），结果显示出相反的趋势。我们布局预测模型的平均均方根 $|\Delta S|$ 从（相对于原始标签的）7.8×10^{-4} 降至 5.6×10^{-4}（逐结构比值中位数为 0.66×）。相比之下，MFT-NNLS 的平均误差从 1.27×10^{-3} 升至 1.50×10^{-3}（中位数为 1.20×），Ansys Genequiv 从 1.39×10^{-3} 升至 1.65×10^{-3}（1.15×），MATLAB RF Toolbox 则从 4.0×10^{-3} 升至 4.1×10^{-3}（1.03×，这一趋势在很大程度上被较高的基线误差所掩盖）。由于逐结构拟合器针对单次电磁求解结果进行优化，当模型保真度接近标签噪声底限时，其极点和留数会部分吸收特定运行所产生的求解器抖动。相反，我们的网络将确定性布局映射为端口哈密顿实现，训练数据涵盖数万个结构，因此在训练过程中能够平均掉不相关的求解器噪声，同时在推理时不受测试标签波动的影响。这种对比凸显了跨布局学习的一项隐式去噪优势：它捕捉的是底层物理响应，而非单次求解中的数值伪影。

五、结论

本文提出了一种端到端宏模型化框架，可将物理布局直接映射为嵌套式、连续频率的无源端口哈密顿宏模型。针对先进封装数据集中的 27 万个 8 端口结构开展实验，并在其中 2.7 万个未见过的测试结构上进行验证，结果证实：与成熟的拟合工具相比，平均误差平均降低 42%（最高降低 65%），无源性违规为零，且宏模型生成时间低于 100 毫秒（加速比为 11，000–39，000×）。

参考文献

[1] D. Kim、J. Park、Y. Oh 和 B. Hwang，“TraceFormer：基于图变换器的印制电路板走线 S 参数预测框架”，载于第 *61* 届 *ACM/IEEE* 设计自动化会议（*DAC*）论文集。美国加利福尼亚州旧金山：ACM，2024 年，第 1–6 页。[2] T. Paik、J. Park、T. Kim、D. H. Jung 和 D. Kim，“PCBFormer：理解真实世界印制电路板走线的三维结构以预测 S 参数”，载于*IEEE/ACM* 计算机辅助设计国际会议（*ICCAD*）论文集。德国慕尼黑：IEEE，2025 年，第 1–8 页。[3] L. Liu、Y. Li、B. Lu、L. Shang 和 F. Yang，“GNN-SP：通过图神经网络快速估算小芯片互连的 S 参数”，*IEEE*元件、封装与制造技术汇刊，第 14 卷，第 10 期，第 1862–1871 页，2024 年 10 月。[4] L. Zhang、J. Juang、Z. Kiguradze、B. Pu、S. Jin、S. Wu、Z. Yang、J. Fan 和 C. Hwang，“使用深度学习快速预测供电网络阻抗”，国际数值建模期刊：电子网络、器件与场，第 35 卷，第 2 期，文章编号 e2956，2022 年。[5] F. Feng、C. Zhang、J. Ma 和 Q.-J. Zhang，“结合神经网络与基于极点-留数的传递函数对微波元件电磁行为进行参数化建模”，*IEEE*微波理论与技术汇刊，第 64 卷，第 1 期，第 60–77 页，2016 年 1 月。[6] S. Grivet-Talocia 和 B. Gustavsen，无源宏模型化：理论与应用，Wiley 微波与光学工程系列。Wiley，2016 年。[7] J. R. Phillips、L. Daniel 和 L. M. Silveira，“用于模型降阶的保证无源平衡变换”，*IEEE*计算机辅助设计集成电路与系统汇刊，第 22 卷，第 8 期，第 1027–1041 页，2003 年 8 月。[8] P. Triverio、S. Grivet-Talocia、M. S. Nakhla、F. G. Canavero 和 R. Achar，“电气互连模型的稳定性、因果性与无源性”，*IEEE*先进封装汇刊，第 30 卷，第 4 期，第 795–808 页，2007 年 11 月。[9] Ansys, Inc.，“Ansys HFSS三维布局帮助，Ansys电磁套件 2024 R1”，https://ansyshelp.ansys.com，2024 年。[10] B. Gustavsen 和 A. Semlyen，“通过向量拟合对频域响应进行有理逼近”，*IEEE*电力输送汇刊，第 14 卷，第 3 期，第 1052–1061 页，1999 年 7 月。[11] B. Gustavsen，“通过扰动留数矩阵特征值快速强制实现 S 参数模型的无源性”，*IEEE*先进封装汇刊，第 33 卷，第 1 期，第 257–265 页，2010 年 2 月。[12] S. Grivet-Talocia，“通过扰动哈密顿矩阵强制实现无源性”，*IEEE*电路与系统汇刊 *I*：常规论文，第 51 卷，第 9 期，第 1755–1769 页，2004 年 9 月。[13] C. P. Coelho、J. R. Phillips 和 L. M. Silveira，“一种通过凸规划生成表格化频率数据的保证无源逼近的方法”，*IEEE*计算机辅助设计集成电路与系统汇刊，第 23 卷，第 2 期，第 293–301 页，2004 年 2 月。[14] P. Benner、P. Goyal 和 P. Van Dooren，“根据频率响应数据识别端口哈密顿系统”，系统与控制快报，第 143 卷，文章编号 104741，2020 年。[15] L. Zeng、J. Sun、X. Wu、D. Niu、T. Wang、Y. Lin、Z. Ye 和 Z. Jin，“G-SpNN：用于神经网络 S 参数建模的 GPU加速无源性强制”，载于第 *62* 届 *ACM/IEEE* 设计自动化会议（*DAC*）论文集。美国加利福尼亚州旧金山：IEEE，2025 年，第 1–7 页。[16] F. Ferranti、L. Knockaert 和 T. Dhaene，“具有保证无源性的基于参数化 S 参数的宏模型化 ”，*IEEE* 微波与无线元件快报，第 19 卷，第 10 期，第 608–610 页，2009 年 10 月。[17] T. Bradde、S. Grivet-Talocia、A. Zanco 和 G. C. Calafiore，“一致稳定且无源的参数化宏模型的数据驱动提取”，*IEEE Access*，第 10 卷，第 15 786–15 804 页，2022 年。[18] P. Schwerdtner 和 M. Schaller，“面向参数化系统的基于结构化优化的模型降阶”，*SIAM*科学计算杂志，第 47 卷，第 1 期，第 A72–A101 页，2025 年。[19] H. M. Torun、A. C. Durgun、K. Aygün 和 M. Swaminathan，“使用神经网络对 S 参数进行因果且无源的参数化”，*IEEE*微波理论与技术汇刊，第 68 卷，第 10 期，第 4290–4304 页，2020 年 10 月。[20] P. Xu、S. Xu、T. Chen、G. Chen、T. Ho 和 B. Yu，“DeepOTF：学习方程约束的电磁行为预测”，*ACM*电子系统设计自动化汇刊，第 29 卷，第 5 期，第 1–22 页，2024 年 9 月，文章编号 80。[21] A. van der Schaft 和 D. Jeltsema，“端口-哈密顿系统理论：入门概述”，系统与控制基础与趋势，第 1 卷，第 2–3 期，第 173–378 页，2014 年。[22] T. Kailath，线性系统。新泽西州恩格尔伍德克利夫斯：Prentice-Hall，1980 年。[23] C. A. Beattie、V. Mehrmann 和 P. Van Dooren，“无源系统的鲁棒端口-哈密顿表示”，*Automatica*，第 100 卷，第 182–186 页，2019 年 2 月。[24] R. Bhatia，正定矩阵，普林斯顿应用数学系列。新泽西州普林斯顿：普林斯顿大学出版社，2007 年。[25] S. Brody、U. Alon 和 E. Yahav，“图注意力网络的注意力程度如何？”，载于国际学习表征会议（*ICLR*）论文集，2022 年。[26] N. Vyas、D. Morwani、R. Zhao、I. Shapira、D. Brandfonbrener、L. Janson 和 S. Kakade，“SOAP：使用 Adam 改进并稳定用于语言建模的 Shampoo”，载于国际学习表征会议（ *ICLR* ）论文集，2025 年。[27] B. Gustavsen，“MFT-NNLS：使用向量拟合和留数扰动，从大型导纳与散射数据集进行无源宏模型化的工具箱”，*IEEE*电力输送汇刊，第 41 卷，第 4 期，第 1652–1662 页，2026 年 8 月。[28] The MathWorks, Inc.，“rationalfit：使用稳定有理函数对象近似数据（RF Toolbox 文档）”，https://www.mathworks. com/help/rf/ref/rationalfit.html，2026 年。[29] Ansys, Inc.，“在命令行运行状态空间拟合（genequiv），电路帮助，Ansys电磁套件 2024 R1”，https://ansyshelp.ansys.com，2024 年。[30] P. C. Opdenacker 和 E. A. Jonckheere，“保持平衡降阶方案的压缩映射及其无穷范数误差界”，*IEEE*电路与系统汇刊，第 35 卷，第 2 期，第 184–189 页，1988 年 2 月。[31] B. C. Moore，“线性系统中的主成分分析：可控性、能观性与模型降阶”，*IEEE*自动控制汇刊，第 26 卷，第 1 期，第 17–32 页，1981 年。