通过插值低秩张量增强非线性系统辨识

Christina Auer，*IEEE*研究生学生会员，Thomas Paireder，*IEEE*研究生学生会员，Oliver Ploder，*IEEE*研究生学生会员，Oliver Lang，*IEEE*会员，以及Mario Huemer，*IEEE*高级会员

arXiv:2302.07505v1 [math.ST] 15 Feb 2023

根据输入和输出数据进行函数逼近，是信号处理领域研究最多的问题之一。该问题已通过各种信号处理和机器学习方法加以解决。尽管张量在众多学科中有着悠久的应用历史，但近年来，基于张量的估计在系统辨识中日益受到关注。本文聚焦于利用插值张量方法解决自适应非线性系统辨识问题。我们提出了三种新方法，将现有的基于张量的估计技术与多维线性插值相结合。为保持较低的复杂度，我们沿用相关思路：算法采用维纳结构或Hammerstein结构，并将张量与广为人知的LMS算法结合。张量的更新基于随机梯度下降。除此之外，对张量更新和LMS采用适当的步长归一化，有助于保证收敛。最后，我们通过多项实验表明，所提出的算法几乎总是明显优于先进水平的方法，同时复杂度更低或相当。

***Index* 术语—插值、LMS、低秩逼近、机器学习、张量分解**

n

I. 引言

IN 近几十年来，科学界一直高度关注系统辨识问题。针对这一问题，基于机器学习的方法获得了极大的普及。机器学习主要与深度学习和神经网络 [1]–[5],相关，但该领域还涵盖许多其他技术，例如随机森林 [6], [7],、支持向量机 [8]–[10], 、核自适应滤波器 [11]–[13]以及基于张量的学习（参见例如 [14]）。

尽管基于张量的方法能够实现与上述其他技术相近的性能 [15], [16]，但由于其较高的计算复杂度，它们长期以来一直被忽视。然而，近年来，基于张量的方法逐渐受到越来越多的关注。基于张量的算法已被用于许多任务，例如人脸识别、乐谱挖掘以及检测社交网络中的团 [14], [17], [18]。此外，它们已经应用于系统辨识 [19]–[21], 、信道辨识 [22],[23], 、语音信号分离 [24], 、雷达发射机定位、通信应用和被动感知 [25],[26]。

复杂度问题已在多种形式和论文中得到探讨。张量列车 [27] 旨在降低整个系统的复杂度和内存占用。为了减少张量所需的维度，会将其拆分成多个较小的张量，并将它们依次级联。在 [28], [29], 中，一种有趣的张量与扩展卡尔曼滤波器的组合得到了研究，旨在提升性能并保持张量的秩较低。此外，在 [30] 中，张量列车与B样条相结合，用以缓解复杂度问题，同时利用B样条在系统辨识任务中的优势。

在 [31],（本文也以此为基础）中，采用了 Scarpiniti 等人 [32] 的方法进行系统辨识，将遵循 Wiener 或 Hammerstein模型 [33] 的系统分解为线性部分和非线性部分。然而，这类方法的一大缺点是问题是连续的。张量本质上是一个巨大的查找表（查找表），而定义张量大小的离散化点可能不足。因此，输出可能并不总是足够准确，进而导致性能大幅下降。

在本研究中，我们还研究了系统辨识，并假设未知系统会随时间变化。我们采用了 [31] 中的自适应张量组合方法。为缓解精度不足的问题，我们将这些经典张量方法与插值相结合。事实上，我们提出了三种新颖的插值张量方法：仅插值张量方法、Wiener系统和Hammerstein系统。后两种方法将张量变体与有限冲激响应（FIR）滤波器相结合，这些滤波器采用广为人知的LMS算法进行训练，用于建模系统的线性部分。插值通过以下方式实现：首先，将张量中与待进行插值的输入样本距离最近的输入样本所对应的元素，存储在一个维度为 $2\times2\times\ldots\times2$ 的独立子张量中。其次，将该子张量的元素视为控制点，并对其进行线性插值，以获得更准确的输出。张量更新采用随机梯度下降（SGD）方法。此外，通过适当的新颖步长归一化确保收敛。最后，我们进行了数值实验，以验证这些研究结果

致谢：谨此感谢奥地利联邦数字与经济事务部、国家研究、技术与发展基金会以及 Christian Doppler 研究协会提供的经费支持。C. Auer、T. Paireder 和 O. Ploder 任职于奥地利林茨约翰内斯·开普勒大学信号处理研究所的 Christian Doppler 数字辅助未来移动通信射频收发器实验室（电子邮箱：christina.auer@jku.at）。O. Lang 和 M. Huemer 任职于奥地利林茨约翰内斯·开普勒大学信号处理研究所

实验验证了插值带来的精度大幅提升。总之，本文的创新贡献如下:

- 插值张量的数学框架
- 采用插值张量方法的三种模型：– 插值分解张量模型– 插值张量-LMS模型– 插值LMS-张量模型，以及
- 步长归一化。

本文的组织结构如下。为完整起见，第二节介绍本文余下部分所需的一些基本符号和基础知识。第三节推导新颖的插值张量模型，第四节则引入一项约束以确保收敛。随后，第五节介绍计算复杂度，第六节展示一些实验结果。最后，第七节总结本文。

II. 基础知识

在物理学和数学中，“张量”一词有不同的含义。本文将张量定义为两个集合之间的多线性映射，这种定义已被广泛采用，也见于 [15], [16]。作为基础，我们首先介绍 [31] 采用的符号，并重述 [31]中著名张量方法最重要的性质和定义。此外，我们还简要概述多维线性插值。这些基础知识将在第三节的推导中用到。

A. 预备知识与符号

张量 $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_M}$ 可以表示为一个 M 维数组，其索引为 $\mathcal{X}(i_1, i_2, i_3、\ldots\ldots、i_M)$。张量 $\mathcal{X} \in \mathbb{R}^{I_1 \times \cdots \times I_M}$ 与向量 $\mathbf{v} \in \mathbb{R}^{I_m \times 1}$ 的乘法可以通过向量的 m 模乘积实现

$$\begin{aligned}&\left(\mathcal{X} \times_m \mathbf{v}^{\mathrm{T}}\right)(i_1, \ldots, i_{m-1}, 1, i_{m+1}, \ldots, i_M)\\ &= \sum_{i_m=1}^{I_m} \mathcal{X}(i_1, \ldots, i_M)\mathbf{v}(i_m).\end{aligned} \tag{1}$$

注意，结果张量的第 m 个维度为 1，这意味着张量的阶已降至 $M-1$。典型多项式分解（CPD）

$$\mathcal{X}(i_1, i_2, \ldots, i_M) = \sum_{r=1}^{R} \prod_{m=1}^{M} \mathbf{A}_m(i_m, r), \tag{2}$$

其中 R 表示秩，可以看作矩阵奇异值分解的推广。

将张量重塑为矩阵称为矩阵化、展开或扁平化。这可以沿所有坐标进行。设 m' $m' \in \{1, \ldots, M\}$ 是任意但固定的数。由此得到的矩阵记为 $\mathcal{X}_{(m')}$，其中索引 (m') 表示张量的扁平化方式。例如，在三维情况下，立方体可以水平切分、沿前方切成薄片，或从侧面切分。无论维度如何，由此得到的矩阵都可以将张量重塑为一个矩阵 $\mathcal{X}_{(m'\ \in)} \in \mathbb{R}^{I_1 \cdots I_{m'-1} I_{m'+1} \cdots I_M \times I_{m'}}$。下面介绍一些常用符号

如 [15], [16], 所示，我们用 $\circledast$ 表示 Hadamard 积，并用 $m \neq m'$ 表示所有矩阵 $\mathbf{A}_m$ 的乘积，其中

$$\circledast_{m \neq m'} \mathbf{A}_m := \mathbf{A}_M \circledast \cdots \circledast \mathbf{A}_{m'+1} \circledast \mathbf{A}_{m'-1} \circledast \cdots \circledast \mathbf{A}_1. \tag{3}$$

B. 张量分解

利用上述所述的分解张量，我们可以通过最小化以下代价函数来研究随机梯度下降算法 [16]

$$J_{\mathrm{SGD}} = \left(\mathcal{Y}(i_1, i_2, \ldots, i_M) - \sum_{r=1}^{R} \prod_{m=1}^{M} \mathbf{A}_m(i_m, r)\right)^2 =: e^2 \tag{4}$$

其中 $\mathcal{Y}$ 是一个包含相应数据点的张量。为求解这个最小化问题，我们需要计算梯度

$$\frac{\partial J_{\mathrm{SGD}}}{\partial \mathbf{A}_{m'}(i_{m'}, :)} = -2e \circledast_{m \neq m'} \mathbf{A}_m(i_m, :). \tag{5}$$

因此可以构造如下更新方程

$$\begin{aligned}&\mathbf{A}_{m',n+1}(i_{m',n+1}, :) =\\ &\qquad \mathbf{A}_{m',n}(i_{m',n}, :) - \mu \left(\frac{\partial J_{\mathrm{SGD}}}{\partial \mathbf{A}_{m'}(i_{m'}, :)}\right)^{\mathrm{T}}\end{aligned} \tag{6}$$

其中 μ 是步长。在可用的数据点中，我们随机选取一个 $\mathcal{Y}$（i_1、i_2、.....、i_M），并且只对那些会受到影响的模型参数执行梯度步骤。由于这里讨论的是更新步骤，因此还需要一个额外的迭代参数 n。为清晰起见，上文的（6）中省略了该参数，在不必要时也会继续省略。稍后在问题描述中，n 还将作为所出现时间信号的时间索引。

C. 多维线性插值

线性插值是一种利用线性多项式，在一组已知离散数据点的范围内构造新数据点的方法。可以将其理解为：利用该离散数据集寻找一个连续函数。我们将多维线性插值称为一种扩展，用于在规则网格上对具有多个变量的函数反复进行线性插值。它近似计算函数f在该网格内某个中间点（$\check{v}_1$、$\check{v}_2$、.....$\check{v}_M$）处的值。待插值的函数 f 在给定点（$v_{1、n}$、$v_{2、n}$、$\ldots v_{M、n}$）处已知，目标是求出任意点（$\check{v}_1$、$\check{v}_2$、.....$\check{v}_M$）处的函数值。考虑单位正方形上的双线性插值（= 个变量的插值，即反复进行线性插值），则有

$$f(\check{v}_1, \check{v}_2) \approx \begin{bmatrix} 1-\check{v}_1 & \check{v}_1 \end{bmatrix} \begin{bmatrix} f(0,0) & f(0,1) \\ f(1,0) & f(1,1) \end{bmatrix} \begin{bmatrix} 1-\check{v}_2 \\ \check{v}_2 \end{bmatrix}. \tag{7}$$

对于具有 M 个维度的多维函数，这将展开为

$$f(\check{v}_1, \ldots, \check{v}_M) \approx \sum_{l_1, \ldots, l_M=0}^{1} f(l_1 \ldots l_M) \prod_{m=1}^{M} \check{\mathbf{v}}_m(l_m), \tag{8}$$

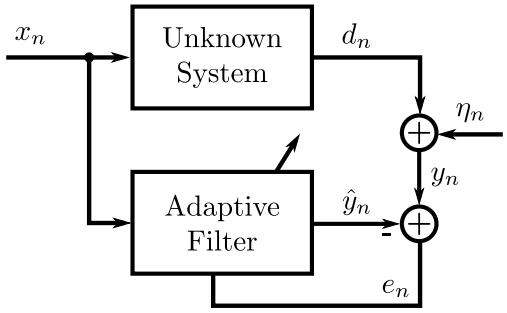


图1：自适应系统辨识的一般设置。通过自适应滤波器对未知系统进行估计。本文中，该系统采用Wiener结构或Hammerstein结构。图片取自 [31]。

其中 $\check{\mathbf{v}}_m = \begin{bmatrix} 1 - \check{} & v_m \\ & v_m \end{bmatrix}\check{}$。请注意，该求和表示多重求和，因此需要对每个 $l_m \in \{0,1\}$ 求和 M 次。

D. 问题描述

本文旨在近似一个未知系统，其一般设置如图1所示。因此，我们对未知系统使用相同的输入信号 x_n，其中 n 表示时间索引，并观测含噪输出 $y_n = d_n + \eta_n$，其中 d_n 为期望信号， η_n 为噪声。目标是最小化误差 $e_n = y_n - \hat{}y_n$，其中$\hat{}y$ 为自适应滤波器的输出，而(·)始终定义该近似结果。由于加性噪声的存在，该误差不可能真正等于零。由于我们假设未知系统可能不会随时间保持静态，因此不会关闭自适应过程，并且自适应滤波器会根据每个观测样本进行更新（即每个时间步执行一次自适应步骤）。需要注意的是，我们希望近似的未知系统也可能具有时间记忆，因此它不仅取决于当前的输入信号 x_n ，还取决于其过去的样本。因此，将 x_n 中最近的 P 个样本按照以下方式存储在抽头延迟线（TDL）中

$$\mathbf{x}_n := \text{tdl}(x_n, \mathbf{x}_{n-1}) := [x_n, \mathbf{x}_{n-1}^{\mathrm{T}}(1:P-1)]^{\mathrm{T}} \tag{9}$$

与 $\mathbf{x}_0 = \mathbf{0}$ 和 $\mathbf{x}_n \in \mathbb{R}^P$。

在本文后续部分，我们将研究六种不同的系统，并使用不同的方法对其进行近似。其中三种未知系统（参见图1）假设遵循 Hammerstein结构（参见图2a），另外三种系统则假设遵循 Wiener结构（参见图2b）。

E. 离散化

分解张量的输入是矩阵 $\mathbf{A}_m$ 的索引。因此，自然只能有 I_m （$\mathbf{A}_m$ 第一维的大小）个不同的输入值，而不能有任意多个。由于输入信号 x_n 通常不局限于 I_m 个不同值，因此必须经过单独的离散化步骤。具体做法是将输入值的最大值与最小值之间的范围划分为若干区间，这些区间也由此决定矩阵 $\mathbf{A}_m$ 的第一维。若最大值和／或最小值未知，则可以定义一个截断点，

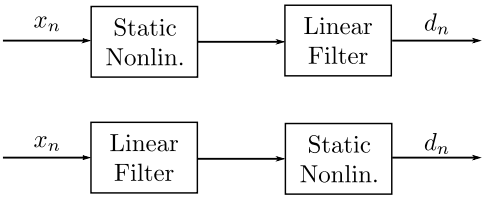


图2：所考虑的未知系统遵循以下结构之一：(a) Hammerstein结构，或 (b) Wiener结构。图片引自[31]。

最大值和／或最小值未知。若 I_m为偶数，则该离散化步骤可用数学形式表示为

$$\text{disc}(x_n) = \left\lfloor \frac{x_n}{\Delta x} \right\rfloor + \frac{I_m}{2} \tag{10}$$

其中， Δx 是离散化间隔。为简化起见，张量的所有 M 维度均采用相同的 Δx 。将不同维度采用不同离散化间隔的情况进行扩展也很简单。向下取整算子 $\lfloor\cdot\rfloor$ 定义如下

$$\lfloor\cdot\rfloor := \max\{k \in \mathbb{Z} \mid k \le x\}. \tag{11}$$

III. 插值张量自适应系统辨识

我们首先介绍基本张量模型。在此基础上，我们介绍三种算法:

- 插值分解张量模型
- 张量-LMS模型，以及
- LMS-张量模型。

A. 基本张量模型

目标是通过对控制点进行线性插值来估计函数值$f(x_1$、.....、$x_M)$，这些控制点是张量$\mathcal{Q}$的元素。该张量的每个条目都对应一个函数值，其中函数的第一维为 x_1，第二维为 x_2，依此类推。

$$\mathcal{Q}(k_1, \ldots, k_M) = f\left(\left(k_1 - \frac{I_1}{2}\right)\Delta x, \ldots\right) \tag{12}$$

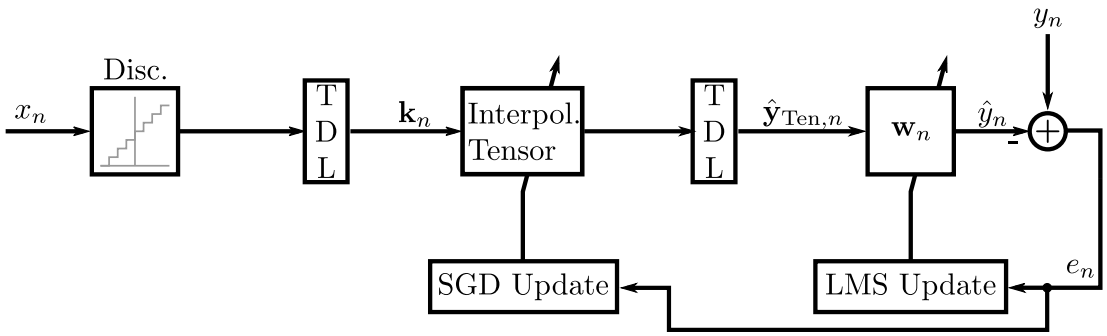
其中，对于所有 $m = 1, \ldots, M$，均有 $k_m = \text{disc}(x_m)$。在这个完整张量中，我们考虑一个子张量

$$\mathcal{Q}_{k_1,\ldots,k_M} \in \mathbb{R}^{2\times 2\times \ldots \times 2} \tag{13}$$

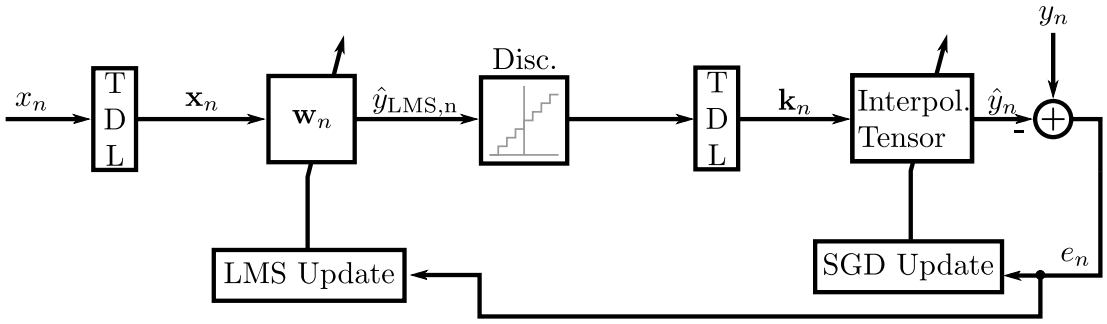
对于每个待估计的函数值，该子张量由式 (13) 所述的点及其相邻点组成。

利用这些知识以及第二节-C 中关于多维线性插值的数学基础，我们现在可以通过对该子张量中的值进行一般插值，估算函数 f 的任意值

$$f(x_1, \ldots, x_M) \approx \mathcal{Q}_{k_1,\ldots,k_M} \times_M \check{\mathbf{u}}_M \ldots \times_1 \check{\mathbf{u}}_1 = \sum_{l_1,\ldots,l_M=0}^{1} \mathcal{Q}_{k_1,\ldots,k_M}(l_1, \ldots, l_M) \prod_{m=1}^{M} \check{\mathbf{u}}_m(l_m) =: \hat{y}_{\text{SGD}} \tag{14}$$



(a) 首先，输入样本经过离散化步骤和抽头延迟线。随后执行插值张量步骤。由于 LMS 的输入需要再次成为向量，因此在此步骤之前要经过一个 TDL。



(b) 在这里，输入样本经过类似的步骤，但 LMS 和张量的顺序互换。输入样本首先由 TDL 汇集，然后输入 LMS。随后对输出进行离散化；在经过插值张量之前，样本还必须通过 TDL。

图 3：用于 Hammerstein 和 Wiener 模型的 (a) 张量最小均方和 (b) LMS-张量架构，摘自 [31]，其中 TDL 表示抽头延迟线。

with

$$\check{\mathbf{u}}_m = \begin{bmatrix} 1-u_m \\ u_m \end{bmatrix} \tag{15}$$

$$u_m = \frac{x_m}{\Delta x} - \left\lfloor \frac{x_m}{\Delta x} \right\rfloor. \tag{16}$$

更新子张量 $\boldsymbol{\mathcal{Q}}_{k_1,\ldots,k_M}$，并由此更新完整张量$\boldsymbol{\mathcal{Q}}$的一种方法是采用随机梯度下降（SGD）。在这里，我们使用平方误差代价函数

$$J = \left(y_n - \hat{y}_{\mathrm{SGD}}\right)^2 := e_n^2. \tag{17}$$

为了解决最小化问题，我们计算导数

$$\frac{\partial J}{\boldsymbol{\mathcal{Q}}_{k_{1,n},\ldots,k_{M,n},n}(l'_1,\ldots,l'_M)} = -2\,e_n \prod_{m=1}^{M} \check{\mathbf{u}}_{m,n}(l'_m) \tag{18}$$
$$:= -2\,e_n\,\boldsymbol{\mathcal{B}}_{k_{1,n},\ldots,k_{M,n},n}(l'_1,\ldots,l'_M),$$

针对 $l'_m \in \{0,1\}$ 的每种组合。利用该导数，我们得到整个子张量的以下更新公式

$$\boldsymbol{\mathcal{Q}}_{k_{1,n},\ldots,k_{M,n},n+1} = \boldsymbol{\mathcal{Q}}_{k_{1,n},\ldots,k_{M,n},n} + 2\mu\,e_n\,\boldsymbol{\mathcal{B}}_{k_{1,n},\ldots,k_{M,n},n}. \tag{19}$$

B. 插值分解张量模型

我们回顾一下常规张量分解 (2)，该分解已在多篇文献中得到研究 [15], [16], [31]

$$\boldsymbol{\mathcal{Q}}(i_1,\ldots,i_M) = \sum_{r=1}^{R}\prod_{m=1}^{M}\mathbf{A}_m(i_m,r). \tag{20}$$

将 (20) 代入 (14)，我们可以使用第 II-C 节讨论的多维线性插值来近似函数值

$$f(x_1,\ldots,x_M) \approx \sum_{l_1,\ldots,l_M=0}^{1}\sum_{r=1}^{R}\prod_{m=1}^{M}\mathbf{A}_m(k_m+l_m,r)\check{\mathbf{u}}_m(l_m). \tag{21}$$

此外，我们可以使用随机梯度下降来更新矩阵 $\mathbf{A}_m$。因此，我们需要代价函数的梯度，其中估计的函数值按式（20）计算。对于所有 $m' \in [1,\ldots,M]$ 和 $l'_{m'} \in \{0,1\}$，可通过计算以下偏梯度得到梯度

$$\frac{\partial J}{\partial \mathbf{A}_{m'}(k_{m'}+l'_{m'},:)} = -2\,e\,\check{\mathbf{u}}_{m'}(l'_{m'}) \tag{22}$$
$$\sum_{l_1,\ldots,l_{m'-1},l_{m'+1},\ldots,l_M=0}^{1} \circledast_{m\neq m'}\mathbf{A}_m(k_m+l_m,:)\check{\mathbf{u}}_m(l_m),$$

其中，$\circledast_{m\neq m'}$ 表示根据式（3）对所有矩阵$\mathbf{A}_m$ 与 $m \neq m'$ 进行 Hadamard 积。

C. 张量最小均方（*TLMS*）

在本小节中，我们假设图1中的未知系统由一个非线性模块和一个线性滤波器组成，符合Hammerstein模型。为估计这个非线性系统，我们采用图3a所示的张量最小均方架构（TLMS）来识别该系统。在此，非线性模块由一个张量建模，张量输出按照

第三节-B中的方法进行插值。张量的输出样本由抽头延迟线收集，随后经过一个线性FIR滤波器，该滤波器用于建模线性部分，其滤波器系数由LMS算法更新。

目标是估计线性滤波器输出端的值，但输入首先经过插值张量分解（参见图3a），从而形成两个前向路径模块。代价函数是瞬时平方误差（如LMS滤波器通常采用的那样）

$$J_{\text{TLMS}} = (y_n - \hat{y}_{\text{LMS},n})^2 =: e_n^2 \tag{23}$$

with

$$\hat{y}_{\text{LMS},n} = \mathbf{w}_n^{\text{T}} \hat{\mathbf{y}}_{\text{Ten},n} = \sum_{p=1}^{P} w_{n,p} \hat{y}_{\text{Ten},n-p+1}, \tag{24}$$

其中 $\hat{\mathbf{y}}_{\text{Ten},\,n}$:=tdl($\hat{y}_{\text{Ten},\,n}$) 表示插值张量输出。插值张量输出表示为

$$\hat{y}_{\text{Ten},n} = \sum_{l_1,\ldots,l_M=0}^{1} \sum_{r=1}^{R} \prod_{m=1}^{M} \mathbf{A}_{m,n}(k_{m,n}+l_m, r) \check{\mathbf{u}}_{m,n}(l_m). \tag{25}$$

因此，我们可以重新表述代价函数

$$\begin{aligned} J_{\text{TLMS}} = &\left(y_n - \sum_{p=1}^{P} w_{n,p} \sum_{l_1,\ldots,l_M=0}^{1} \sum_{r=1}^{R} \right. \\ &\left. \prod_{m=1}^{M} \mathbf{A}_{m,n-p+1}(k_{m,n-p+1}+l_m, r) \check{\mathbf{u}}_{m,n-p+1}(l_m) \right)^2 \\ &=: (y_n - s_{n-p+1}) = e_n^2. \end{aligned} \tag{26}$$

为了能够计算张量和LMS权重的更新量，我们需要计算反向路径中的导数。我们首先进行张量求导

$$\begin{aligned} &\frac{\partial J_{\text{TLMS}}}{\partial \mathbf{A}_{m'}} = -2\, e_n \sum_{p=0}^{P} w_p \\ &\left(\begin{pmatrix} \mathbf{0}_{k_{m',n-p+1}-1\times R} \\ \frac{\partial s_{m',n-p+1}}{\partial \mathbf{A}_{m'}(k_{m',n-p+1},:)} \\ \mathbf{0}_{I_{m'}-k_{m',n-p+1}\times R} \end{pmatrix} + \begin{pmatrix} \mathbf{0}_{k_{m',n-p+1}\times R} \\ \frac{\partial s_{m',n-p+1}}{\partial \mathbf{A}_{m'}(k_{m',n-p+1}+1,:)} \\ \mathbf{0}_{I_{m'}-k_{m',n-p+1}-1\times R} \end{pmatrix} \right) \\ &=: -2\, e_n\, \mathbf{S}_{m',n} \end{aligned} \tag{27}$$

其中 $\mathbf{0}_{k_{m',\,n-p+1}-1\times R} \in \mathbb{R}^{k_{m'},\,{}_{n-p+1}-1\times R}$ 表示一个所有元素均为零的矩阵，并且

$$\begin{aligned} &\frac{\partial s_{m',n-p+1}}{\partial \mathbf{A}_{m'}(k_{m',n-p+1}+l'_{m'},:)} = \\ &\check{\mathbf{u}}_{m',n-p+1}(l'_{m'}) \sum_{l_1,\ldots,l_{m'-1},l_{m'+1},\ldots,l_M=0}^{1} \\ &\circledast_{m\neq m'} \mathbf{A}_{m,n-p+1}(k_{m,n-p+1}+l_m,:) \check{\mathbf{u}}_{m,n-p+1}(l_m). \end{aligned} \tag{28}$$

LMS的反向路径非常直接

$$\frac{\partial J_{\text{TLMS}}}{\partial \mathbf{w}_n} = -2\, e_n\, \hat{\mathbf{y}}_{\text{Ten},n}. \tag{29}$$

注意，由于 $\mathbf{w}_n$ 是列向量，且我们采用分子布局，因此 $\frac{\partial J_{\text{TLMS}}}{\partial \mathbf{w}_n}$ 是行向量。利用这些方程，我们可以计算分解张量模型的更新量

$$\mathbf{A}_{m',n+1} = \mathbf{A}_{m',n} + 2\mu_2 e_n\, \mathbf{S}_{m',n}. \tag{30}$$

LMS 的更新通常表示为 [34] 并给出如下

$$\mathbf{w}_{n+1} = \mathbf{w}_n + 2\mu_1\, e_n\, \hat{\mathbf{y}}_{\text{Ten},n}. \tag{31}$$

D. LMS-张量（*LMST*）

为了解决这样的问题：图1中的未知系统可表示为一个线性项后接一个非线性，我们提出一种插值LMS张量（LMST）架构，参见图3b。线性滤波器仍由FIR滤波器建模，其滤波器系数通过LMS算法更新；而非线性则由一个张量建模，其输出按照第三节-B所述进行插值。代价函数为

$$J_{\text{LMST}} = (y_n - \hat{y}_{\text{Ten},n})^2 =: e_n^2. \tag{32}$$

张量的前向路径以及由此得到的插值输出很简单，表示为

$$\hat{y}_{\text{Ten},n} = \sum_{l_1,\ldots,l_M=0}^{1} \sum_{r=1}^{R} \prod_{m=1}^{M} \mathbf{A}_{m,n}(k_{m,n}+l_m, r) \check{\mathbf{u}}_{m,n}(l_m). \tag{33}$$

其中，在此情况下，$k_{m,\,n} = \text{disc}(\hat{y}_{\text{LMS},\,n})$ 是 LMS算法的输出，而 disc(·) 描述根据（10）执行的离散化步骤。此外，LMS 的前向路径很简单，表示为

$$\hat{y}_{\text{LMS},n} = \mathbf{w}_n^{\text{T}} \mathbf{x}_n = \sum_{p=1}^{P} w_{n,p} x_{n-p+1}. \tag{34}$$

因此，代价函数可概括为

$$\begin{aligned} &J_{\text{LMST}} = \\ &\left(y_n - \sum_{l_1,\ldots,l_M=0}^{1} \sum_{r=1}^{R} \prod_{m=1}^{M} \mathbf{A}_{m,n}(k_{m,n}+l_m, r) \check{\mathbf{u}}_{m,n}(l_m) \right)^2. \end{aligned} \tag{35}$$

从 LMS 的反向路径开始，我们计算关于权重向量的导数；应用链式法则后，得到

$$\frac{\partial J_{\text{LMST}}}{\partial \mathbf{w}_n} = \sum_{m'=1}^{M} \frac{\partial J_{\text{LMST}}}{\partial \check{\mathbf{u}}_{m',n}} \frac{\partial \check{\mathbf{u}}_{m',n}}{\partial \hat{y}_{\text{LMS},n-m'+1}} \frac{\partial \hat{y}_{\text{LMS},n-m'+1}}{\partial \mathbf{w}_n}. \tag{36}$$

分别来看各项，我们计算关于 $\check{\mathbf{u}}_{m',\,n}$ 的导数，得到

$$\begin{aligned} \frac{\partial J_{\text{LMST}}}{\partial \check{\mathbf{u}}_{m',n}} = &-2\, e_n \sum_{r=1}^{R} \mathbf{A}_{m',n}(k_{m',n} : k_{m',n}+1, r) \\ &\sum_{l_1,\ldots,l_{m'-1},l_{m'+1},\ldots,l_M=0}^{1} \prod_{\substack{m=1 \\ m\neq m'}}^{M} \\ &\mathbf{A}_{m,n}(k_{m,n}+l_m, r) \check{\mathbf{u}}_{m,n}(l_m) \end{aligned} \tag{37}$$

接着计算关于 $\hat{y}_{\text{LMS},\ n-m+1}$的导数

$$\frac{\partial \check{\mathbf{u}}_{m',n}}{\partial \hat{y}_{\text{LMS},n-m'+1}} = \begin{pmatrix} -\frac{1}{\Delta x} \\ \frac{1}{\Delta x} \end{pmatrix}, \tag{38}$$

最后但同样重要的是，关于权重

$$\frac{\partial \hat{y}_{\text{LMS},n-m'+1}}{\partial \mathbf{w}_n} = \mathbf{x}^{\text{T}}_{n-m'+1}. \tag{39}$$

汇总这些导数，可得

$$\frac{\partial J_{\text{LMST}}}{\partial \mathbf{w}_n} = -2\, e_n \sum_{m'=1}^{M} \Delta a_{m',n} \sum_{r=1}^{R} \mathbf{v}_{m',n}(r)\mathbf{x}^{\text{T}}_{n-m'+1} \tag{40}$$

with

$$\Delta a_{m',n} = \sum_{r=1}^{R} \left(\frac{1}{\Delta x}\mathbf{A}_{m',n}(k_{m',n}+1,r) - \frac{1}{\Delta x}\mathbf{A}_{m',n}(k_{m',n},r) \right) \tag{41}$$

and

$$\mathbf{v}_{m',n}(r) = \sum_{l_1,\ldots,l_{m'-1},l_{m'+1},\ldots,l_M=0}^{1} \prod_{\substack{m=1 \\ m\neq m'}}^{M} \mathbf{A}_{m,n}(k_{m,n}+l_m,r)\check{\mathbf{u}}_{m,n}(l_m). \tag{42}$$

对于反向路径，需要求出 J_{LMST} 关于 $\mathbf{A}_{m',\ n}(k_{m',\ n}+l'_{m',}\ :)$ 的导数，可推导如下

$$\frac{\partial J_{\text{LMST}}}{\partial \mathbf{A}_{m',n}(k_{m',n}+l'_{m'},:)} = -2\, e_n\, \check{\mathbf{u}}_{m',n}(l'_{m'})\mathbf{v}_{m',n} \tag{43}$$

对于所有 $m' \in \{1, M\}$ 和 $l_{m'} \in \{0,1\}$。最后，利用 (40) 和 (43)，我们可以计算滤波器系数和分解张量的更新量，如下所示

$$\mathbf{w}_{n+1} = \mathbf{w}_n + 2\mu_1\, e_n \sum_{m'=1}^{M} \Delta a_{m',n} \sum_{r=1}^{R} \mathbf{v}_{m',n}(r)\mathbf{x}^{\text{T}}_{n-m'+1} \tag{44}$$

and

$$\mathbf{A}_{m',n+1}(k_{m',n+1}+l'_{m'},:) = \mathbf{A}_{m',n}(k_{m',n}+l'_{m'},:) + 2\mu_2\, e_n\, \check{\mathbf{u}}_{m',n}(l'_{m'})\mathbf{v}_{m',n}. \tag{45}$$

请注意，此处的 (45) 必须针对每个 $m'\ \in\{1,\ldots,M\}$进行更新。

IV. 步长归一化

为提高所提出方法的稳定性，我们根据上一小节的内容，推导出插值张量-LMS 和 LMS-张量的更新参数μ 的归一化形式。为此，我们考虑将先验误差e_{n+1} 近似为一阶泰勒级数（见 [35]）。

A. 张量-*LMS*归一化

本小节研究张量最小均方的归一化。因此，我们用一阶泰勒展开近似新的误差

$$e_{n+1} \approx e_n + \sum_{r=1}^{R} \frac{\partial e_n}{\partial \mathbf{A}_{m'}(:,r)} \Delta \mathbf{A}_{m'}(:,r) \tag{46}$$

with

$$e_n = y_n - \mathbf{w}_n^{\text{T}} \hat{\mathbf{y}}_{\text{Ten},n}. \tag{47}$$

由（30）可得

$$\Delta \mathbf{A}_{m'}(:,r) = 2\mu_2 e_n \mathbf{S}_{m',n}(:,r) \tag{48}$$

and

$$\frac{\partial e_n}{\partial \mathbf{A}_{m'}(:,r)} = \mathbf{S}^{\text{T}}_{m',n}(:,r). \tag{49}$$

由（46）、（48）和（49）可得

$$\begin{aligned} e_{n+1} &\approx e_n - 2\mu_2 e_n \sum_{r=1}^{R} \mathbf{S}^{\text{T}}_{m',n}(:,r)\mathbf{S}_{m',n}(:,r) \\ &= e_n - 2\mu_2 e_n \sum_{r=1}^{R}\sum_{i=1}^{I_{m'}} |\mathbf{S}_{m',n}(i,r)|^2 \\ &= e_n - 2\mu_2 e_n \left\|\mathbf{S}_{m',n}\right\|_{\text{F}}^2, \end{aligned} \tag{50}$$

其中 $\|\cdot\|_{\text{F}}$ 表示弗罗贝尼乌斯范数，因此，

$$e_{n+1} \approx \left(1 - 2\mu_2 \left\|\mathbf{S}_{m',n}\right\|_{\text{F}}^2\right) e_n. \tag{51}$$

基于泰勒展开，若步长满足以下条件 [35]，则可提高算法的自适应性能

$$\left|1 - 2\mu_2 \left\|\mathbf{S}_{m',n}\right\|_{\text{F}}^2\right| < 1. \tag{52}$$

因此，可得步长的以下界限 μ_2

$$0 < \mu_2 < \frac{1}{\left\|\mathbf{S}_{m',n}\right\|_{\text{F}}^2}. \tag{53}$$

请注意，该界限必须对每个 $m'\ \in\{1,\ldots,M\}$都成立。可以通过在（30）中将 μ_2替换为时变步长来实现这一界限

$$\mu_{2,n} = \frac{\mu_2}{\delta + \left\|\mathbf{S}_{m',n}\right\|_{\text{F}}^2}, \qquad 0 < \mu_2 < 1. \tag{54}$$

较小的正则化值 $\delta > 0$ 有助于避免弗罗贝尼乌斯范数较小时出现数值问题。

B. LMS-张量归一化

与张量-LMS 情形类似，我们针对 LMS-张量情形推导泰勒展开

$$e_{n+1} \approx e_n + \frac{\partial e_n}{\partial \mathbf{w}_n} \Delta \mathbf{w}_n^{\text{T}}. \tag{55}$$

这可以通过与推导（44）相同的方法完成，其中

$$\frac{\partial e_n}{\partial \mathbf{w}_n} = -\sum_{m'=1}^{M} \Delta a_{m',n} \sum_{r=1}^{R} \mathbf{v}_{m',n}(r)\mathbf{x}^{\text{T}}_{n-m'+1}, \tag{56}$$

and

$$\Delta \mathbf{w}_n^{\mathrm{T}} = 2\mu_1 \, e_n \sum_{m'=1}^{M} \Delta a_{m',n} \sum_{r=1}^{R} \mathbf{v}_{m',n}(r) \mathbf{x}_{n-m'+1}. \quad (57)$$

因此，我们得到

$$\begin{aligned} e_{n+1} &\approx e_n - 2\mu_1 \, e_n \overbrace{\sum_{m'=1}^{M} \Delta a_{m',n} \sum_{r=1}^{R} \mathbf{v}_{m',n}(r) \mathbf{x}_{n-m'+1}^{\mathrm{T}}}^{=\mathbf{d}_n^{\mathrm{T}}} \\ &\quad \underbrace{\sum_{m'=1}^{M} \Delta a_{m',n} \sum_{r=1}^{R} \mathbf{v}_{m',n}(r) \mathbf{x}_{n-m'+1}}_{=\mathbf{d}_n} \\ &= \left(1 - 2\mu_1 \left\| \mathbf{d}_n \right\|_2^2 \right) e_n \, . \end{aligned} \quad (58)$$

按照与之前相同的推理，步长 μ_1 应满足以下界限

$$0 < \mu_1 < \frac{1}{\left\| \mathbf{d}_n \right\|_2^2}. \quad (59)$$

可通过将式 (44) 中的 μ_1 替换为时变步长来实现该界限

$$\mu_{1,n} = \frac{\mu_1}{\delta + \left\| \mathbf{d}_n \right\|_2^2}, \qquad 0 < \mu_1 < 1, \quad (60)$$

其中， δ 是一个较小的正则化值。

V. 计算复杂度

本节从加法、乘法和除法的角度比较本文所涉及的所有方法的复杂度。对于 LMS 算法，由表 I 可知，其更新需要 $2P+1$ 次乘法和 $2P$ 次加法。要计算经典张量在正向路径和反向路径中所需的运算次数，我们考察了第二节中的公式。对于我们提出的插值张量模型，则仔细考察了第三节推导的公式以及第四节中的步长归一化。为简便起见，我们为每个维度选取相同的离散化大小，因此

$$l_0 = l_1 = \cdots = l_m. \quad (61)$$

对于式 (21) 中的仅插值张量方法的前向路径（$M-1$），需要进行R 次乘法和 $R-1$ 次加法。此外，式 (22) 中给出的插值随机梯度下降更新的反向路径更为复杂，其总结见表 I。插值 TLMS 和 LMST 算法的复杂度也可通过考察第三节中的相应公式，以类似方式推导，并以一般形式列于表 I。显然，无论是经典张量方法还是插值方法，秩 R 和维度 M 对每个样本所需的运算次数影响最大。经典张量方法的计算复杂度随维度 M线性增长，而插值张量方法的计算复杂度则随 M呈指数增长。

VI. 性能仿真

我们通过性能仿真结束对插值张量的研究，并将其与文献 [31]提出的经典张量方法进行比较。所有所考察算法的总体目标都是最小化期望信号 $d_n^{(l)}$与其复现信号 $\hat{}y_n^{(l)}$之间的误差，其中 n 表示时间，l 表示运行次数（即第 l次重复）。因此，我们用于比较的指标是归一化均方误差（NMSE）

$$\mathrm{NMSE}_{\mathrm{dB}} = 10 \log_{10} \left(\frac{1}{L} \sum_{l=1}^{L} \frac{\left(d_n^{(l)} - \hat{y}_n^{(l)} \right)^2}{\left(d_n^{(l)} \right)^2} \right), \quad (62)$$

其中， L 是给定实验的总运行次数。请注意，在 d_n 中未包含测量噪声，参见图1。

为了验证所提出的插值张量架构，我们进行了多项实验。本研究中的实验1–4与 [31], [36] 中进行的实验完全对应。为求完整起见，下面再次介绍这些实验。我们评估了前文讨论的所有方法，即仅插值张量、插值 TLMS 和插值 LMST 的性能。各项设置（步长、秩等）均通过经验选取，以使所有算法达到最佳性能。对于所有实验，所有自适应算法的权重和抽头延迟线均初始化为零，张量则初始化为方差为 0.01 的标准正态分布随机值。实验中使用的信噪比（SNR）为 10 dB。在示例1和2中，我们希望突出所考虑算法的自适应特性。因此，在仿真时间过半后，我们改变未知系统的线性部分。这种情况可能是由系统内的温度漂移造成的，也可能只是因为运行期间切换到了其他线性滤波器。

A. 实验1

第一个实验聚焦于识别一个Hammerstein模型，即系统由静态非线性模块和其后的线性部分组成。输入信号 x_n 按照以下关系生成

$$x_n = a x_{n-1} + \sqrt{1-a^2} \nu_n \, , \quad (63)$$

其中， ν_n 是均值为零、方差为单位值的白高斯噪声，$0 \leq a < 1$ 中参数 a 决定相邻样本之间的相关性。关于 a 的选择，详情请参见 [31], [36]。系统的第一个模块由以下形式的静态非线性描述

$$y_n^{\mathrm{nl}} = \frac{2x_n}{1 + \left| x_n \right|^2} \, . \quad (64)$$

系统的这一部分模拟卫星通信中的功率放大器 (PA) 的饱和行为 [37], 参见 [31] 中的第一个实验。第二个模块是一个未知的线性FIR滤波器 $\mathbf{w}_0 \in \mathbb{R}^7$ ，用于模拟传输路径的一部分。在本实验中，我们在运行时间过半后改变这个线性

表 I：根据所考虑算法的参数，估计（前向）步骤和更新（反向）步骤中每个样本所需的算术运算次数。本文研究了已有的经典仅张量、TLMS 和 LMST 方法，以及新推导出的仅插值张量、插值 TLMS 和插值 LMST 方法。参数 P、R、 M 和 I_m 分别表示线性滤波器的长度、秩、张量维数和离散化大小。

算法		乘法	Add.	Div.
LMS	前向路径	P	$P-1$	–
	反向路径	$2P+1$	$2P$	1
仅张量	前向路径	$(M-1)R$	$R-1$	–
	反向路径	$MR(1+R(M-1)+I_m)$	$MR(1+I_m)$	M
intpol 仅张量	前向路径	$2^M MR$	$2^M MR-1$	–
	反向路径	$2M+RM\left(1+2^{M-1}(2M-3)+I_m\right)$	$MR(2^{M-1}+I_m)$	M
TLMS	前向路径	$P+R(M-1)$	$P+R-2$	–
	反向路径	$MR\left(P(M-1)+I_m\right)+2P(M+1)+1$	$2P+MRI_m(P+1)$	$1+M$
LMST	前向路径	$P+R(M-1)$	$P+R-2$	–
	反向路径	$1+2P+M\left(P+1+R(2M-2+I_m)\right)$	$M\left(R(3+I_m)+P-1\right)+P$	$1+M$
intpol TLMS	前向路径	$2^M RM+P$	$2^M RM+P-2$	–
	反向路径	$MRP\left(3+2^{M-1}(2M-3)\right)+I_m R(M+1)+M^2+2P$	$I_m R(1+M)+P(2^{2M-1}+2^{M-1}(1-R)+2)-M$	$1+M$
插值 LMST	前向路径	$2^M RM+P$	$2^M RM+P-2$	–
	反向路径	$P(M+2)+2(M+1)+MR\left(2+2^M(2M-3)+I_m\right)$	$2P+MR(2^{M-1}3+I_m+1)-1$	$1+M$

表 I：对于研究的六种不同实验，所应用的每种算法都需要不同的参数。此处汇总了实验中使用的经典张量方法和插值张量方法的参数值。

实验	张量				LMS		插值张量				插值LMS		仅张量				intpol 仅张量			
	μ_2	R	M	I	μ_1	P	μ_2	R	M	I	μ_1	P	μ	R	M	I	μ	R	M	I
1	0.009	1	1	50	0.009	7	0.01	1	1	10	0.01	7	0.05	50	7	25	0.1	10	3	10
2	0.1	1	1	50	0.0075	5	0.01	1	1	10	0.001	5	0.01	50	5	23	0.1	10	3	10
3	0.008	1	3	50	0.005	5	0.01	10	2	16	0.01	5	0.025	100	7	25	0.4	20	3	10
4	0.05	10	2	16	0.002	7	0.01	10	2	16	0.01	7	0.05	200	8	25	0.1	20	3	32
5	0.1	30	3	50	0.02	3	0.1	16	3	20	0.8	3	0.09	40	3	30	0.1	40	3	20
6	0.4	10	4	40	0.001	4	0.5	16	2	30	0.001	2	0.4	20	3	20	0.4	20	3	10 1

系统的系数。$\mathbf{w}_0$ 系数最初设定为

$\begin{pmatrix}0.9 & -0.6 & 0.3 & -0.15 & 0.1 & -0.025 & 0.005\end{pmatrix}^{\mathrm{T}}$ as in [36] and later by $\begin{pmatrix}0.6 & -0.4 & 0.25 & -0.15 & 0.1 & -0.05 & 0.001\end{pmatrix}^{\mathrm{T}}$, cf. [31].

将仅插值张量方法以及由于 Hammerstein结构而采用的插值 TLMS架构，与经典仅张量方法和 TLMS方法进行比较。图 4a 展示了四种算法在归一化均方误差方面的性能。可以看出，所提出的两种插值变体均显著优于各自的经典对应方法。尽管两个 TLMS 版本的秩和维度都等于 1，但由于插值情形下只需较低的离散化程度，其复杂度更低（参见表 II 和表 III）。对于纯张量方法，经典方法所需的维度和秩远高于其插值版本。尽管如此，几乎所有方法都能快速响应，并按照预期适应先前提到的线性子系统变化，参见 [31]。例外的是经典仅张量方法，它不仅需要远远最多的算术运算，性能也不如其他方法。该方法的性能不够理想，也无法恰当地适应系统变化。变化后，其表现甚至更差。

B. 实验*2*

第二个实验采用 Wiener模型。此处的总体设置与前一个实验使用相同的非线性，不同之处在于非线性位于线性滤波器之后，而该滤波器仅包含 5 个抽头；详情请参见 [31]。因此，本例采用 LMST 结构。图 4b 展示了各算法的性能。可以看出，所提出的两种方法同样优于各自的经典对应方法。与前一个实验类似，LMST 变体中张量的秩 R 和维度 M 均设为 1，而经典仅张量方法所需的阶数和秩远高于插值版本。与前一个例子一样，除经典仅张量方法外，所有算法都能按照预期适应系统变化，并在变化后取得相近的性能。

C. 实验*3*

第三个实验再次采用Wiener模型设置，使用5抽头线性滤波器，但我们选择了不同的非线性

$$y_n^{\mathrm{nl}} = \sin^2(x_n) + \sin^3(x_{n-1}) + \sin^4(x_{n-2})\,. \tag{65}$$

与实验二中的简单情形不同，时刻 n 的非线性输出取决于该块当前及前两个输入值。因此，该非线性具有记忆。输入信号 x_n 遵循与

式（63）中实验1的关系。对于此示例，图4c也表明，所提出的方法优于经典张量解。插值 LMST 的表现最佳，其次是纯插值张量解。从复杂度来看，插值 LMST 比其经典对应方法更复杂。相比之下，纯插值张量解所需的复杂度低于经典仅张量变体（参见表III）。

D. 实验*4*

第四个实验再次采用 Hammerstein模型，其设置与实验3相同，但由7个抽头组成的线性部分与非线性部分互换。图4d中的评估结果表明，就归一化均方误差和收敛速度（通过自适应开始时归一化均方误差曲线的斜率衡量）而言，仅插值张量显著优于经典纯张量解，而插值组合版本仅略优于经典版本。在复杂度方面，LMST算法之间的差异很小。经典方法需要更多加法，而插值方法的乘法次数更多。

E. 实验*5*

第五个实验与第一个实验十分相似，同样采用Hammerstein结构。此外，该系统的设计灵感来自一个特定的移动通信问题，而非卫星通信。其非线性特性受到发射端谐波问题 [38] 的启发，同时也模拟了功率放大器的非理想特性。它建模为

$$y_n^{\mathrm{nl}} = x_n^2\,, \tag{66}$$

后接一个长度为 3 的线性滤波器。输入信号不满足关系式 (63)，而是方差为 1 的白高斯噪声。图 4e 展示了本次仿真的结果。所提出的方法显著优于经典张量方法。虽然插值 TLMS 变体的表现略逊于仅插值张量方案，但其优势在于所需的算术运算更少。

F. 实验*6*

第六个实验与采用 Wiener 结构的第二个实验相似。该实验同样受到一个特定移动通信问题的启发，即二阶互调失真（IMD2）问题 [39]。线性滤波器有 3 个抽头，非线性部分的形式如下

$$y_n^{\mathrm{nl}} = |x_n|^2\,. \tag{67}$$

输入数据再次选为方差为 1 的白高斯噪声。图 4f 中的评估结果表明，两种插值方法均优于各自的经典对应方法，且所有方法都具有较快的收敛速度。

G. 比较

每项实验所需的运算次数列于表III。这表明，TLMS/LMST版本的复杂度始终低于纯张量对应方法。此外可以看出，在所有仿真中，插值纯张量方法所需的运算次数均少于经典版本；而对于Wiener和Hammerstein方法，情况则因示例而异。有时，插值方法所需的运算次数更少。在所有情况下，就最终归一化均方误差而言，插值版本均优于经典方法。只有在实验4中，TLMS方法的表现几乎与插值 TLMS方法相当。由于大多数所考虑的方法具有相似的自适应表现，因此较难评估收敛速度。可以明确指出两个例外：经典纯张量方法存在收敛问题；实验6中的经典LMST方法也表现出较差的自适应行为。总体而言，与输入样本具有式（63）所示结构的情况相比（参见实验1至4），经典TLMS/LMST变体似乎更难处理正态分布输入。总之，我们提出的插值 TLMS/LMST方法是低成本架构，其性能可以优于先进水平方法及其经典对应方法。

VII. 结论

本文针对一般系统辨识问题，提出了三种将先进水平的张量方法与插值相结合的新型自适应学习方案。这三种方案分别是仅插值张量方法、Hammerstein 系统学习算法和 Wiener 系统学习算法。后两种方法基于广为人知的 LMS 算法。多个实验通过归一化均方误差（ NMSE ）展示了所提方法的有效性。此外，为确保稳定性，本文还推导出了学习率选择的约束条件。另外，复杂度分析表明，在全部六个实验中，所提出的仅插值张量方法所需运算量均少于经典方法，而插值 TLMS/LMST方法的复杂度也几乎相同。

参考文献

[1] W. Liu、Z. Wang、X. Liu、N. Zeng、Y. Liu 和 F. E. Alsaadi，《深度神经网络架构及其应用综述》，*Neurocomputing*，第234卷，第11–26页，2017年4月。[2] K. Burse、R. N. Yadav 和 S. C. Shrivastava，《使用神经网络进行信道均衡：综述》，*IEEE*系统、人类与控制论汇刊*C*辑：应用与评论，第40卷，第3期，第352–357页，2010年11月。[3] O. Ploder、O. Lang、T. Paireder 和 M. Huemer，《一种基于自适应机器学习的方法，用于消除4G/5G收发器中的二阶互调失真》，载于90th 车辆技术会议 - VTC 2019秋季会议论文集。美国檀香山：IEEE，2019年9月，第1–7页。[4] A. Balatsoukas-Stimming，《使用神经网络对带内全双工无线电进行非线性数字自干扰消除》，载于 19th 无线通信信号处理进展国际研讨会 - SPAWC 2018论文集，希腊卡拉马塔，2018年6月，第1–5页。

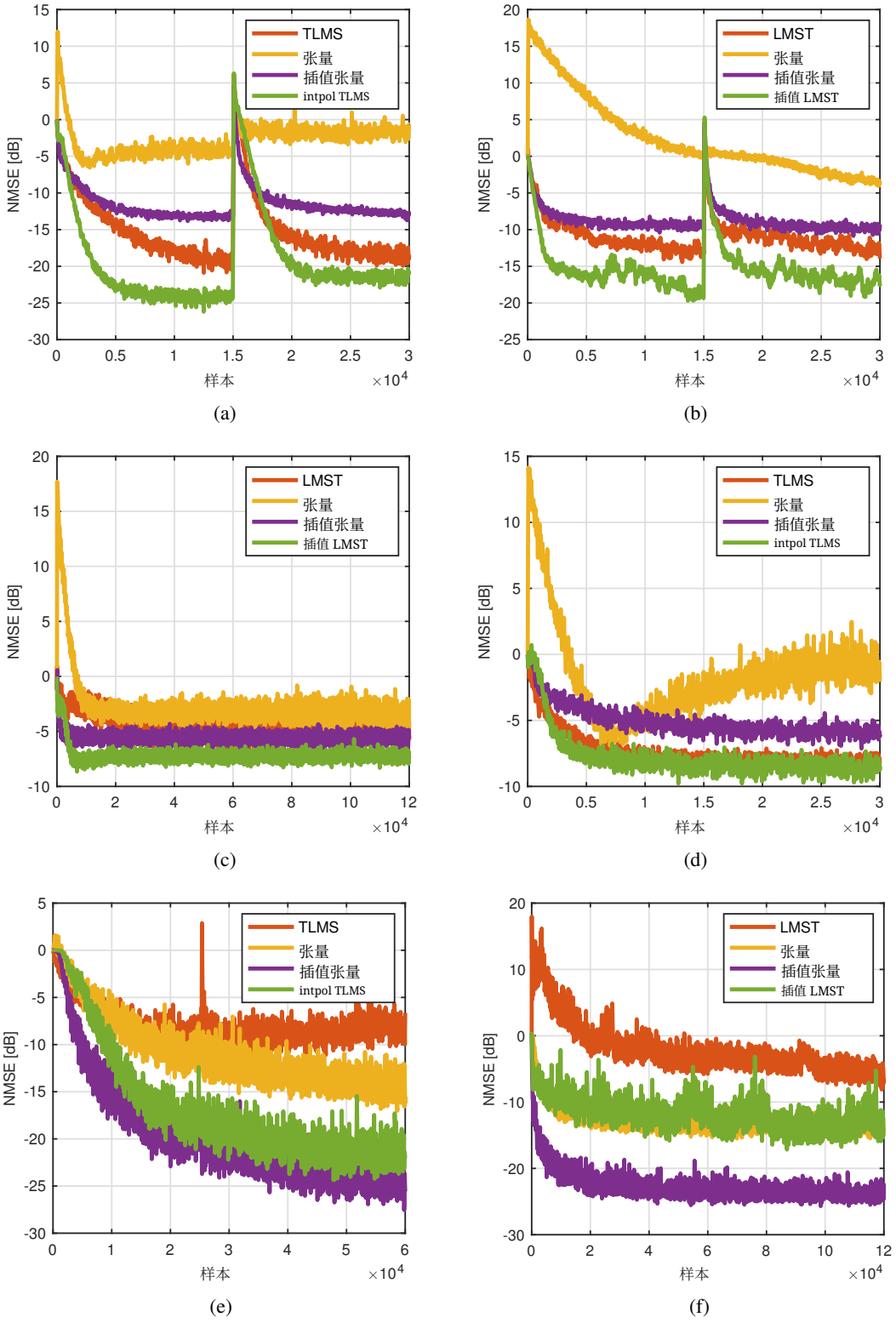


图 4：所有所考虑算法的归一化均方误差收敛比较：(a) Hammerstein情形；(b) Wiener情形；(c) 含记忆的非线性的 Wiener 情形；(d) 非线性依赖过去样本的 Hammerstein 情形；(e) 受移动通信启发的 Hammerstein 情形；(f) 受移动通信启发的 Wiener模型

表III：所有实验中所考虑算法的乘法、加法和除法复杂度。

实验	运算	仅张量	TLMS/LMST	插值 仅张量	intpol TLMS/LMST
1	乘法	114400	86	936	58
	Add.	9149	420	659	54
	Div.	7	2	3	2
2	乘法	56200	72	1866	36
	Add.	6049	66	1319	28
	Div.	5	2	3	2
3	乘法	438800	198	1866	551
	Add.	18299	180	1319	512
	Div.	7	4	3	3
4	乘法	2283000	520	3186	1285
	Add.	41799	2589	2639	467
	Div.	8	3	3	3
5	乘法	12200	1308	4926	3842
	Add.	2559	3863	3839	1328
	Div.	3	4	3	4
6	乘法	3100	1053	1866	1296
	Add.	679	1076	1319	1251
	Div.	3	3	3	3

[5] O. Ploder、C. Motz、T. Paireder 和 M. Huemer，“一种用于消除未来蜂窝射频收发器中二阶互调失真的神经网络方法”，载于 53rd 阿西洛马信号、系统与计算机会议 - ACSSC 2019论文集，美国加利福尼亚州太平洋丛林，2019年11月，第1144–1148页。[6] T. Hastie、R. Tibshirani 和 J. Friedman，统计学习要素，施普林格统计学丛书。美国纽约州纽约：施普林格纽约公司，2001年。[7] M. A. H. Shaikh 和 K. Barbé，“采用随机森林对 Wiener-Hammerstein 系统进行初始估计”，载于国际仪器与测量技术会议 - I2MTC 2019论文集，新西兰奥克兰，2019年5月，第1–6页。[8] C. Auer、K. Kostoglou、T. Paireder、O. Ploder 和 M. Huemer，“用于移动通信收发器自干扰消除的支持向量机”，载于第91届车辆技术会议 - VTC Spring 2020论文集，比利时安特卫普，2020年，第1–6页。[9] A. J. S mola 和 B. Schölkopf，“支持向量回归教程”，《统计与计算》，第14卷，第3期，第199–222页，2004年8月。 [在线]。获取地址：https://doi.org/10.1023/B: STCO.0000035301.49549.88[10] C. Cortes 和 V. Vapnik，“支持向量网络”，机器学习，第20卷，第3期，第273–297页，1995年9月。 [在线]。获取地址：https://doi.org/10.1007/BF00994018。[11] C. Auer、A. Gebhard、C. Motz、T. Paireder、O. Ploder、R. Sunil Kanumalli、A. Melzer、O. Lang 和 M. Huemer，“核自适应滤波器：移动通信收发器自干扰消除的万灵药？”，载于《计算机科学讲义》(LNCS)：计算机辅助系统理论 - EUROCAST 2019论文集，西班牙大加那利岛拉斯帕尔马斯，2019年2月，第36–43页。[12] W. Liu、J. Principe 和 S. Haykin，核自适应滤波。威利，2010年。[13] R. A. Kennedy 和 P. Sadeghi，信号处理中的希尔伯特空间方法。英国剑桥：剑桥大学出版社，2013年3月。[14] A. Cichocki、D. Mandic、L. De Lathauwer、G. Zhou、Q. Zhao、C. Caiafa 和 H. A. PHAN，“信号处理应用中的张量分解：从二路到多路成分分析”，《*IEEE*信号处理杂志》，第32卷，第2期，第145–163页，2015年3月，会议名称：《IEEE信号处理杂志》。[15] N. D. Sidiropoulos、L. De Lathauwer、X. Fu、K. Huang、E. E. Papalexakis 和 C. Faloutsos，“用于信号处理和机器学习的张量分解”，*IEEE*信号处理汇刊，第65卷，第13期，第3551–3582页，2017年7月，arXiv：1607.01668。 [在线]。获取地址：http://arxiv.org/abs/1607.01668。[16] N. Kargas 和 N. D. Sidiropoulos，“基于张量补全的非线性系统辨识”，*arXiv:1906.05746 [cs, stat]*，2019年12月，arXiv：1906.05746。 [在线]。获取地址：http://arxiv.org/abs/1906.05746。[17] M. A. O. Vasilescu 和 D. Terzopoulos，“图像集合的多线性分析：张量脸”，载于 7th 欧洲会议《计算机视觉 - ECCV 2002》，A. Heyden、G. Sparr、M. Nielsen 和 P. Johansen 编，丹麦哥本哈根，2002年5月，第 447–460 页。[18] E. E. Papalexakis、C. Faloutsos 和 N. D. Sidiropoulos，“Parcube：稀疏可并行张量分解”，载于《欧洲机器学习与数据库知识发现联合会议 - ECML PKDD 2012》论文集，英国布里斯托尔，2012年9月，第 521–536 页。[19] M. Boussé、O. Debals 和 L. De Lathauwer，“利用分割实现基于张量的大规模盲系统辨识”，*IEEE*信号处理汇刊，第 65 卷，第 21 期，第 5770–5784 页，2017年12月。[20] G. Favier 和 A. Y. Kibangou，“基于张量的系统辨识方法”，自动控制与计算机工程科学与技术国际期刊 - *STA 2008*，第 3 卷，第 1 期，第 840–869 页，2008年7月。[21] M. Sørensen 和 L. De Lathauwer，“具有块托普利兹结构的张量分解及其在信号处理中的应用”，载于第 45 届《阿西洛马信号、系统与计算机会议 - ACSSC 2011》论文集，美国加利福尼亚州太平洋丛林，2011年11月，第 454–458 页。[22] C. A. Fernandes、G. Favier 和 J. C. M. Mota，“利用张量分解和预编码实现多用户非线性信道盲辨识”，*Signal Processing*，第 89 卷，第 12 期，第 2644–2656 页，2009年6月。[23] A. Kibangou 和 G. Favier，“利用 Paratuck-2 张量分解对维纳–哈默斯坦通信信道进行盲联合辨识与均衡”，载于 15th 《欧洲信号处理会议 - EUSIPCO 2007》论文集。波兰波兹南：IEEE，2007年9月，第 1516–1520 页。[24] D. Nion、K. N. Mokios、N. D. Sidiropoulos 和 A. Potamianos，“卷积语音混合信号的基于 PARAFAC 的批处理与自适应盲分离”，*IEEE*音频、语音与语言处理汇刊，第 18 卷，第 6 期，第 1193–1207 页，2010年8月。[25] D. Nion 和 N. D. Sidiropoulos，“MIMO 雷达信号处理中的张量代数与多维谐波估计”，*IEEE*信号处理汇刊，第 58 卷，第 11 期，第 5693–5705 页，2010年7月。[26] N. D. Sidiropoulos、R. Bro 和 G. B. Giannakis，“传感器阵列处理中的平行因子分析”，*IEEE*信号处理汇刊，第 48 卷，第 8 期，第 2377–2388 页，2000年8月。[27] C. Y. Ko、K. Batselier、W. Yu 和 N. Wong，“利用张量列车实现快速、精确的张量补全：一种系统辨识方法”，*CoRR*，第 abs/1804.06128 卷，2018年。 [在线]。获取地址：http://arxiv.org/abs/1804.06128。[28] K. Batselier、C.-Y. Ko 和 N. Wong，“用于 MIMO Volterra 系统辨识的低秩张量网络扩展卡尔曼滤波”，载于 58th 《决策与控制会议 - CDC 2019》论文集，法国尼斯，2019年12月，第 7148–7153 页。[29] K. Batselier 和 N. Wong，“张量网络卡尔曼滤波器的矩阵输出扩展及其在 MIMO Volterra 系统辨识中的应用”，*arXiv:1708.05156 [cs]*，2017年8月，arXiv：1708.05156。[在线]。获取地址：http://arxiv.org/abs/1708.05156。

[30] R. Karagoz 和 K. Batselier，“采用正则化张量网络 B 样条的非线性系统辨识”，*arXiv:2003.07594 [cs, eess, math]*, 2020年3月，arXiv: 2003.07594。 [在线]。获取地址：http://arxiv.org/abs/2003.07594。[31] C. Auer、O. Ploder、T. Paireder、P. Kovács、O. Lang 和 M. Huemer，“基于低秩张量分解的自适应系统辨识”，*IEEE Access*，第 9 卷，第 139 028–139 042 页，2021年。[32] M. Scarpiniti、D. Comminiello、R. Parisi 和 A. Uncini，“非线性样条自适应滤波”，*Signal Processing*，第 93 卷，第 4 期，第 772–783 页，2013年4月。[33] N. Wiener，随机理论中的非线性问题。MIT出版社和 John Wiley & Sons, Inc.，1958年。[34] S. M. Kay，统计信号处理基础，A. V. Oppenheim 编。Prentice Hall，美国上萨德尔河，1993年。[35] A. I. Hanna 和 D. P. Mandic，“一种针对复值非线性自适应滤波器的全自适应归一化非线性梯度下降算法”，*IEEE*信号处理汇刊，第 51 卷，第 10 期，第 2540–2549 页，2003年10月。[36] M. Scarpiniti、D. Comminiello、R. Parisi 和 A. Uncini，“用于非线性系统辨识的新型级联样条架构”，*IEEE*电路与系统汇刊第一辑：正规论文，第 62 卷，第 7 期，第 1825–1835 页，2015年7月。[37] T. Panicker、V. Mathews 和 G. Sicuranza，“自适应并联级联截断沃尔泰拉滤波器”，*IEEE*信号处理汇刊，第 46 卷，第 10 期，第 2664–2673 页，1998年。[38] C. Auer、T. Paireder 和 M. Huemer，“用于发射机引起的自干扰消除的核递归最小二乘算法”，载于 93rd《车辆技术会议 - VTC 2021春季会议》论文集，芬兰赫尔辛基，2021年4月，第 1–5 页。[39] C. Auer、T. Paireder、O. Lang 和 M. Huemer，“基于核递归最小二乘的二阶互调失真消除”，载于 54th《阿西洛马信号、系统与计算机会议 - ACSSC 2020》论文集，美国加利福尼亚州太平洋丛林，2020年11月，第 1–5 页。



Christina Auer (GSM’18) 于1987年出生于奥地利林茨。2005年至2012年，她在林茨约翰·开普勒大学学习应用数学。她的硕士论文课题是图像去噪，其中部分内容是在加州大学洛杉矶分校完成的。毕业后，她加入了LCM（林茨机电一体化中心），从事研发工作，尤其是用于定位的卡尔曼滤波。自2017年9月起，她成为林茨约翰·开普勒大学信号处理研究所的成员。目前，她正在攻读博士，作为“面向未来移动通信的数字辅助射频收发器”Christian Doppler实验室的一员。她的研究工作主要聚焦于利用核自适应滤波和机器学习进行自干扰消除。



Thomas Paireder （GSM’19）于2016年和2018年分别获得林茨约翰·开普勒大学（JKU，奥地利林茨）的信息电子学荣誉学士学位和电子与信息技术荣誉硕士学位。在硕士论文中，他实现了一套用于激光超声信号的实时处理系统，该系统应用于无损检测。他目前正在 JKU 信号处理研究所攻读博士学位，研究方向是通过自适应信号处理实现接收机干扰消除处理方法。



Oliver Ploder （S’18）于1992年出生于奥地利格拉茨。2012年至2016年，他在林茨约翰·开普勒大学学习，并获得信息电子学学士学位。2018年，他获得加泰罗尼亚理工大学（巴塞罗那理工）的电信工程硕士学位，专攻无线通信。他的硕士论文与加州大学洛杉矶分校（UCLA）的网络化与嵌入式系统实验室（NESL）合作完成，研究主题是信息物理系统中的安全状态估计系统（CPS）。自2018年9月起，他成为约翰内斯·开普勒大学信号处理研究所的成员，并在那里攻读博士学位，研究重点是利用机器学习对LTE和LTE-A射频收发器系统进行接收机干扰消除。



Oliver Lang （S’14–M’17）1987年出生于奥地利舍尔丁。他在维也纳工业大学学习电气工程，并于2011年获得学士学位。接下来的两年里，Oliver Lang 在维也纳工业大学学习微电子学，并以优异成绩完成学业。他的硕士论文课题是扫描微波显微镜模型的开发与分析。2014年至2018年，他担任约翰内斯·开普勒大学信号处理研究所的大学助理林茨，并于2018年获得博士学位。他的博士论文题目是《估计理论与自适应滤波中的知识辅助方法》，在论文中，他为特殊应用设计了若干有趣的估计器和自适应滤波器。2018年至2019年，他在林茨的 DICE GmbH 工作，该公司是英飞凌奥地利有限公司的子公司。自2019年3月起，他担任信号处理研究所的博士后大学助理。



Mario Huemer （M’00–SM’07）于1996年和1999年分别获得奥地利林茨约翰内斯·开普勒大学（JKU）的工程师文凭和技术博士学位。在工业界和学术界任职后，他于2004年至2007年担任德国埃朗根-纽伦堡大学副教授，并于2007年至2013年担任奥地利克拉根福大学正教授。自2013年9月起，Mario Huemer担任林茨约翰内斯·开普勒大学信号处理研究所所长及正教授；自2017年起，他还共同领导“Christian 面向未来移动通信的数字辅助射频收发器多普勒实验室”。他的研究重点包括统计与自适应信号处理、信号处理架构，以及应用于信息与通信工程、通信与雷达射频收发器、传感器和生物医学信号处理的混合信号处理。在这些领域，他已发表270多篇科学论文。2009年至2015年，他担任《国际电子与通信杂志》（AEU）编委会成员；2017年至2019年，他担任IEEE信号处理快报副编辑。Mario Huemer曾分别获得德国信息技术学会（ITG）和奥地利信息与通信技术学会（GIT）的博士论文奖、奥地利自然科学红衣主教因尼策尔奖，以及德国ITG奖。