arXiv:2301.01521v1 [physics.atom-ph] 4 Jan 2023

人工神经网络：计算正电子湮没寿命谱参数的有效工具

M. Pietrow*[1] 和 A. Miaskowski[2]

[1]物理研究所，玛丽·居里-斯克沃多夫斯卡大学，M. Curie-Skłodowskiej 广场 1 号，20-031 卢布林，波兰[2]生产工程学院，生命科学大学，Akademicka 街 13 号，20-950 卢布林，波兰

2023年1月5日

摘要

本文以固相烷烃为例，介绍了采用多层感知器回归器模型预测正电子湮没寿命谱参数的方法。与常用方法进行比较后发现，计算结果具有良好的一致性。所提出的方法可作为一种快速、准确的替代工具，用于一般情况下正电子湮没寿命谱的分解。本文还讨论了这一新方法的优缺点。

1 引言

正电子湮没寿命谱学（PALS）是一种有用的实验方法，利用正电子研究各种材料，尤其是固态材料中的结构细节 [1]。该方法基于正电子的湮没过程；正电子在样品中的寿命和湮没强度取决于材料在纳米尺度上的一些性质，包括局部电子密度、束缚电子能量，以及样品中自由体积的密度和尺寸。

根据材料的不同，除了直接湮没过程外，正电子还可与电子形成一种称为正电子素（Ps）的亚稳原子态。正电子素有两种自旋态，分别称为仲正电子素和正电子素，其性质各异（尤其是在真空中的寿命相差三个数量级）[2]。正电子素要在物质中形成，必须满足若干条件，其中之一是存在尺寸足够大的自由体积。对于这类材料，正电子素在材料科学中极为有用，因为其寿命与自由体积的尺寸相关 [3]。根据样品的结构，可能存在多种正电子素组分，它们以各自特征寿命发生湮没。所有这些群体都会对实验测得的正电子湮没谱产生贡献。正电子湮没寿命谱需要在测量后的处理中进行分解，从而计算特定正电子种类的寿命以及这些过程的相对振幅（即所谓的谱反演问题） [4]。

许多用于数据处理的算法都需要假设正电子衰变具有指数特性，同时还需要固定分解过程中所用组分的数量。例如，一种适用软件 LT 程序 [5] 或 PALSfit [6], 所采用的方法，是将 PALS 实验光谱拟合为给定数量的指数函数之和，通常还要与（多重）高斯仪器分辨曲线进行卷积。

本文使用的正电子湮没寿命谱是在正烷烃（n-烷烃）中测得的。正烷烃是最简单的有机分子，其碳原子形成分子的直链，并由氢原子饱和。分子中含有不同数量 n 碳原子的正烷烃构成一个同系物系列，其通用化学式为 C_nH_{2n+2} （Cn 为缩写）。烷烃在固相中形成分子晶体，其中排列成列的细长分子彼此由称为层间隙的空隙隔开。正电子素可以形成

*通讯作者：marek.pietrow@umcs.pl

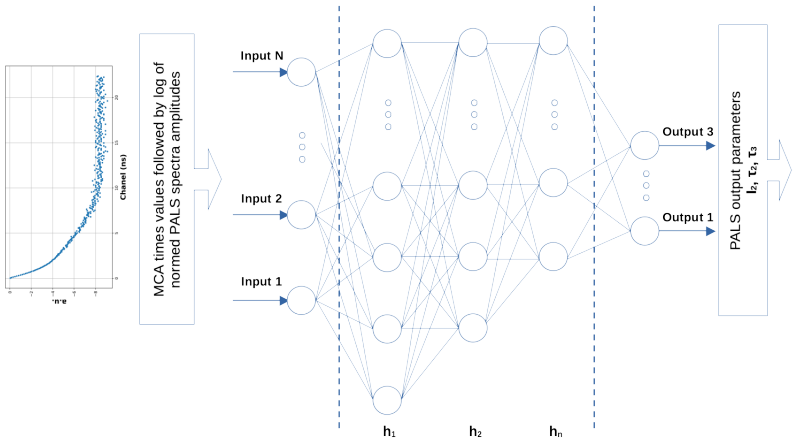


图1：所采用的多层感知机（MLP）示意图。来自多道分析器连续通道的正电子湮没寿命谱数据，以连续输入神经元 In_i 的振幅形式传入。h_i 表示第 i 个隐藏层中的神经元，而 Out_i 表示用于返回所选正电子湮没寿命谱分解参数的输出神经元。

于由层间隙以及温度变化引起的构象改变所产生的空间构成的自由体积中 [3]。利用正电子湮没寿命谱技术，可以根据寿命确定这些自由体积的大小，两者之间的关系由 Tao-Eldrup公式或其修正式给出 [7]。根据我们此前使用正电子湮没寿命谱技术对烷烃进行的分析，假设仲正电子素和 正交-与 仲-正电子素各只有一个群体时，谱分解结果最佳；而 正交与 仲的强度比固定为 3/1。

遗传算法或人工神经网络等机器学习工具已用于正电子科学的多个领域，执行数值计算 [8, 9, 8, 10, 11, 12]。它们也被用于从正电子湮没寿命谱中反演寿命和强度 [13, 14, 15, 16]。可能是由于提出用于正电子湮没寿命谱分解的神经网络算法时，硬件计算能力较低，且正电子湮没寿命谱仪的时间分辨率较低，因此这些计算中使用的大多数光谱都是由软件模拟生成的，而非直接测量所得。出于同样的原因，当时采用的神经网络架构无法像如今开发的算法那样灵活地调整参数。此外，此后也没有提出一种流程，能够对每通道时间常数不同的已采集光谱执行相同的计算。因此，首选的谱分解软件仍基于非线性拟合算法，无法同时根据多谱集合得出结果。

在此，我们提出一种基于多层感知机（MLP）模型分析正电子湮没寿命谱的方法，该模型是机器学习工具之一 [17]。该模型假设网络由相互连接的神经元组成，并将其分为输入层（In_i）、隐藏神经元层（h_i^k）和输出神经元层（Out_i），其中 i遍历给定层中的神经元，k 标示隐藏层的编号。所用网络的示意图如图 1 所示。输入神经元和输出神经元的数量，分别由输入到网络的独立输入数据量，以及在给定问题中被定义为计算结果的数据量决定。隐藏层的数量及这些层中神经元的数量通过实验设定，以优化网络，使其给出所需结果。出于技术原因，每一层（输出层除外）都连接一个偏置神经元 [18]。该方法首先进行学习过程：将已知输出神经元结果的数据输入输入神经元。在此过程中，算法会调整相互连接的神经元对之间的权重系数，使多层感知机的输出尽可能接近期望结果。经过一定次数的训练迭代后，多层感知机便完成训练。若多层感知机的学习结果令人满意，就可以用它计算从未用于训练过程的输入数据所对应的输出。

多层感知机类型的网络可用于解决分类和回归问题。对于第一类问题，多层感知机需要将输出参数的值划分为彼此明确区分的类别。这些所谓的 标签 始终可以用一组离散数值进行参数化。本文所述问题属于回归问题（MLPR），其中每个输出神经元的输出值以一组连续数值表示。因此，输出可能包含接近学习过程中出现的数值，但不一定与其完全相同。MLPR 的内部算法使我们可以将学习过程视为寻找输入参数准连续输出函数的过程。在本研究中，以正电子湮没寿命谱数据作为感知机的输入值，利用 MLPR 解决回归问题，从而求得输出端的关键正电子湮没寿命谱参数。

2 方法

我们使用*scikit-learn* 库来估算烷烃的正电子湮没寿命谱参数 [19]。本研究采用了多层感知机回归器类（称为*MLPRegressor*），它属于监督神经网络模型之一。该类的输出是一组连续值，并使用平方误差作为损失函数。该模型使用属于拟牛顿法的 Broyden–Fletcher–Goldfarb–Shanno算法（L-BFGS） [20],优化平方误差。下文将介绍 MLPRegressor 的一些关键参数。这些参数的取值需要进行调优，尤其是 alpha超参数，它通过惩罚幅值较大的权重来帮助避免过拟合。MLPRegressor 的完整参数列表见 [21]。

在此处的学习过程中，我们使用了模拟谱仪多年来采集的若干种烷烃（C_6 – C_{40}范围内）在多个温度（-142℃ – 100℃）下测得的光谱。无论具体实验的目的如何，也无论作为样品使用的烷烃链有多长，使用 LT 程序 [5] 开始分析光谱时所作的初始假设都相同。每次测量所用的样品均以类似方式制备，即样品经过脱气，冷却或加热速率也相同。每种情况下，恒温测量至少持续一小时，由此产生数十万次湮没事件（各次实验所用放射源的强度相近）。有些实验中，温度按阶梯方式改变，但每条光谱均在恒定温度下采集。这里最重要的是，每次实验后分析光谱时都采用相同的一般假设。特别是在分解这些光谱时，我们使用 LT，并假定时间分辨曲线可用单高斯曲线近似。每次均假定 Kapton包封中的湮没过程占 10%（即所谓的源修正）。此外，始终假定仲正电子素和正电子素各只有一个组分，而它们的强度比固定为 3/1（实验步骤详见 [22]）。

根据这些假设，每个光谱都被分解为三条指数曲线，并计算了以下正电子亚群对应的强度（I）和寿命（τ）：自由正电子湮没（I_2, τ_2）、仲正电子素 （I_1, τ_1）和正电子素（I_3, τ_3）[1]。由此建立的数据库包含 7973 条正电子湮没寿命谱，其中约 75% 用于神经网络训练过程，其余部分用作测试集，以检查训练所得网络结果的准确度。

输入神经元的数量由记录正电子湮没寿命谱的正电子湮没寿命谱仪中多道分析器（MCA）模块的通道数决定。此外，在该模型中，输出神经元的数量与正电子湮没寿命谱参数的数量有关，这些参数旨在预测出来，以便进一步研究样品中的物理过程。使用 LT 等常用程序对正电子湮没寿命谱进行分解，可以确定给定光谱所有假定组分对应的 (I,τ) 对。然而，进一步分析通常并不需要所有这些参数。此外，其中一些参数彼此相关。例如，对于本文讨论的烷烃正电子湮没寿命谱，通常假定该光谱由来自上述三类正电子群体的事件构成（τ_1 – τ_3, I_1 – I_3 参数）。然而，从实际角度来看，研究样品中的物理过程和结构时，只使用 τ_2, I_2, τ_3 和 I_3 。此外，在这种情况下， I_i 彼此相关，并满足以下关系： $I_1+I_2+I_3=100\%$[2] 和 $I_3/I_1=3$。因此，实际上，被视为输出的参数

[1]索引的编号与 τ的长度有关。索引值递增对应寿命长度增加。[2]Kapton 中的湮没已预先扣除。

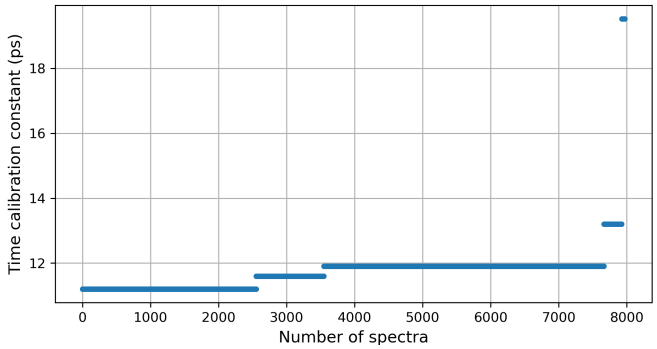


图2：在所示计算中用作数据集的、具有给定每道时间常数 Δ （纵轴）的光谱数量（横轴）。

MLPR 的参数只有 I_2, τ_2 和 τ_3。据此，我们在建模中仅设置了三个输出神经元，用于接收这三个参数的值。

3 输入和输出数据的准备

在正电子湮没寿命谱测量过程中，每道时间常数（Δ）会因能谱仪的内部特性和设置而变化。此处使用的大部分数据是在 Δ=11.9 ps 条件下采集的；不过，也有一些光谱是在 Δ=11.2 ps、13.2 ps、11.6 ps 和 19.5 ps 条件下测得的（图2）。因此，神经元需要将 PALS 振幅样本编码为时间尺度上的数据，而非与道号相关的数据。为此，除了能谱振幅，回归器还必须学习与这些振幅对应的时间值。因此，一半的神经元接收光谱中连续道对应的时间值，另一半则接收这些通道的振幅值。此处采用的回归方法的优点在于，即使测试光谱是在训练过程中从未出现过的极端时间序列条件下测得，也能够对其进行测试。

该方法要求为每个光谱正确设定一个共同的零时刻。为此，使用能谱峰左侧斜坡上的原始数据（以及其右侧少量数据点）对分辨曲线进行插值，并假定其呈高斯形式。该峰的位置定义了光谱的零时刻。单高斯插值符合以往 LT分析的假设。以此确定所有光谱的共同起始位置后，根据测量该光谱时的 Δ，对每个光谱中峰右侧各通道的时间值进行重新校准。最后，为进行进一步分析，我们对每个光谱都取其峰右侧相同数量的 N 个连续道（图 3 中的数据点 p_i ）。图 3 所示的 δ 参数表示峰右侧第一个数据点与计算得到的峰时间位置之间的距离（以时间单位计）。用于进一步分析的数量 N 是通过实验确定的。最后，作为多层感知机回归输入的光谱数据是 N 个数据点 p_i 及其对应的两个值：给定通道中重新校准后的计数（见下文）和重新校准后的湮没时间。

然后，为了尽量减少误差，在应用之前对原始输入数据进行了变换。每条原始光谱储存在多道分析器的8192个通道中。首先，从右侧第一个通道开始，直到光谱最大值处（见图3中的p_1 ），从原始光谱中取出2k个通道。这意味着光谱在约25 ns的采集时间处截断（具体数值会因给定光谱的 Δ而略有不同）。其次，为了平滑随机波动，在大多数情况下对数据进行了平滑处理。平滑处理的一个例子是对五个连续道取平均。在这种情况下，每条光谱中的样本数从原来的2k个通道缩减至400个。由于In神经元传递一对数值——时间（t部分）和振幅（A部分）的信息，因此本例中设置了800个输入神经元，将数据送入多层感知机回归。第三，为了统一输入数据值的范围，将PALS振幅集合按振幅最大值归一化，然后取对数。经过这些变换， A部分数据的数值范围为 [-9,0]——见图4。此外，为了调整 t区域内的数值范围，将时间值除以-2.5。结果，所有传输的数据

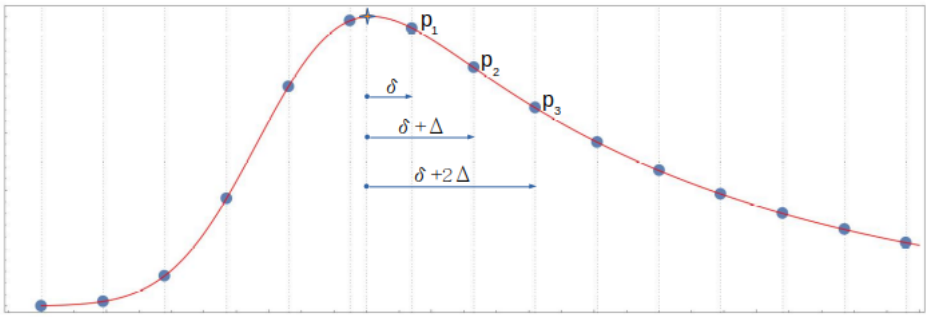


图 3：正电子湮没寿命谱峰区的示意图。圆点表示保存在多道分析器通道中的（t，log（A））数据对，而星号表示假设仪器分布函数呈高斯形状时计算出的峰位置。只有星号右侧的点（p_1、p_2，.....）才作为数据输入多层感知机回归。δ 参数表示计算出的峰与第一个数据点之间的时间间隔，而 Δ 表示两个数据点之间的时间间隔。

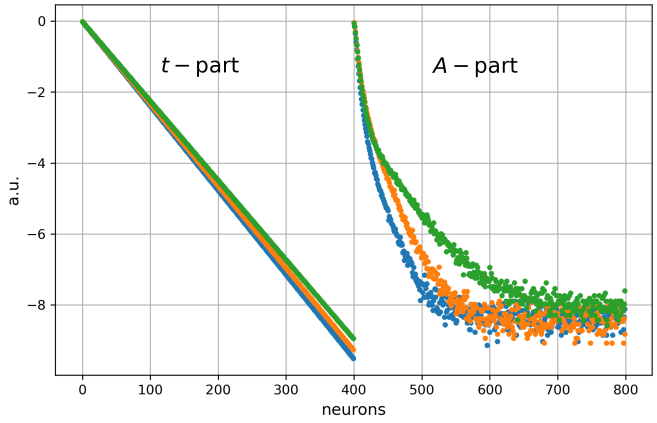


图 4：输入神经元的输入数据可分为两个子集：t部分是数据点 p_1、p_2、.....的时间值集合（见图 3），A部分则对其归一化振幅的对数函数进行编码。在特殊情况下，这些数据在用于多层感知机回归之前会经过平滑或压缩处理。

传递给In神经元的数据都处于 [-10,0]范围内。

此外，我们对输出神经元的原始值进行了一些变换，使每个神经元的值都缩放到相同范围。最初，第一个输出神经元对应于 I_2，其原始值范围通常为几十（以 % 为单位）。第二个神经元传递与τ_2 相关的信息，其原始值通常约为 0.1（以 ns 为单位）；而与第三个神经元 τ_3 相关的值原本约为 1（以 ns 为单位）。为了使所有输出神经元上的数值处于相同量级，最终输入这些神经元的数据分别为 [I_2/10、τ_2·10、τ_3]。

训练网络的验收标准是为该回归器定义的评分验证函数的最佳值

$$\mathcal{S} = 1 - \frac{\sum_N (\mathcal{O}_{\text{true}} - \mathcal{O}_{\text{pred}})^2}{\sum_N \mathcal{O}_{\text{true}}^2}, \tag{1}$$

其中，$\mathcal{O}_{\text{true}}$, $\mathcal{O}_{\text{pred}}$ 分别表示结果的期望值（已知值）和计算值（预测值）[19, 21]。学习集和测试集的 $\mathcal{S}$分别计算。此处的 N 表示训练集或测试集中的光谱数量。$\mathcal{S}$ 的最佳值为 ≈1。

4 结果

本次计算所用的 MLPRegressor 需要设置一些影响学习能力和学习过程速度的关键参数 [21] 。我们进行了若干测试，尝试优化这些参数。通过表1所示的设置，我们获得了最佳结果。这些参数的名称及其含义

Tab表 1：用于生成最终 MPLR 结果的 MLPRegressor 参数值 .

参数	值
隐藏层大小 _	7×150
激活函数 =	*relu*
求解器 =	*L-BFGS*
阿尔法 =	0.01
学习_率 =	逆缩放
幂_t =	0.5
max_迭代次数 =	5e+9
随机 _状态 =	None
容差 =	0.0001
warm_起始值 =	True
max_函数 =	15000

表中所示的技术参数与常规说明中定义的参数相同 [21]。确定多层感知机回归的关键参数（尤其是求解器）后，我们通过改变隐藏层数量、每层神经元数量（hidden_layer_sizes 参数）以及阿尔法参数，对网络的可靠性进行了测试。表 2 展示了部分结果示例。对于这些网络，我们分别列出了训练 $\langle \mathcal{S}_{tr} \rangle$ 集和测试 $\langle \mathcal{S}_{te} \rangle$ 集的平均验证得分参数及其变化 $\delta\mathcal{S}$。平均值是对相同网络进行十次训练过程所得结果计算得出的，这些网络的区别在于初始随机权重不同。我们没有发现任何规则能够确定连续隐藏层中应设置的神经元数量比例（尤其是神经元数量应在连续层中按比例递减这一点）。这里展示的几个初步示例表明，增加隐藏层数量和每层神经元数量都能提高结果的准确度。然而，表格最后两行显示，进一步增加这些参数并不能带来更好的结果。最终，我们选用了结果接近最佳的网络（表中以粗体标出 $\langle \mathcal{S}_{tr} \rangle$ ）。我们检查了该网络，发现将训练迭代次数（max_iter 参数）增加到约 $5{\cdot}10^9$ 以上并未带来改善 $\langle \mathcal{S} \rangle$。

表2：评估 所选部分多层感知机回归参数取值的得分。 $\mathcal{S}$ 数值为采用随机初 th
始化神经元权重 的10次运行的平均值。近似最优的参数组合列于 $\mathcal{S}$ 以粗体标 d.

隐藏_层_大小	最大迭代次数_	阿尔法	$\langle \mathcal{S}_{tr} \rangle$	$\delta\mathcal{S}_{tr}$	$\langle \mathcal{S}_{te} \rangle$	$\delta\mathcal{S}_{te}$
$30 \times 25 \times 15$	10^6	0.7	0.950	0.003	0.942	0.004
3×100	10^8	0.7	0.969	0.005	0.965	0.008
3×100	10^8	0.1	0.974	0.005	0.968	0.008
3×100	$5{\cdot}\ 10^8$	0.1	0.975	0.003	0.976	0.003
4×100	10^8	0.1	0.978	0.004	0.975	0.006
$500 \times 400 \times 300 \times 200$	$5{\cdot}10^9$	0.01	0.977	0.003	0.974	0.007
7×150	$5{\cdot}10^8$	0.01	**0.985**	0.002	0.975	0.013
$500 \times 500 \times 400 \times 400 \times$ $\times 300 \times 300 \times 200 \times 200$	$5{\cdot}10^9$	0.01	0.978	0.005	0.977	0.008
8×500	$5{\cdot}10^9$	0.01	0.982	0.004	0.980	0.005

对于最终测试的若干个网络，我们检查了神经元间权重幅值的光谱。预计与平均值范围相差显著的权重可能会影响结果的稳定性。在这种情况下，权重值的范围似乎相当狭窄。如图5所示，所选网络的权重幅值量级（权重指数）介于 10^{-5} 至 10^0, 之间，而这些子集中的相对案例数呈指数变化。狭窄取值范围之外没有数值

这表明，MLPRegressor算法本身会对最终权重进行自清理。

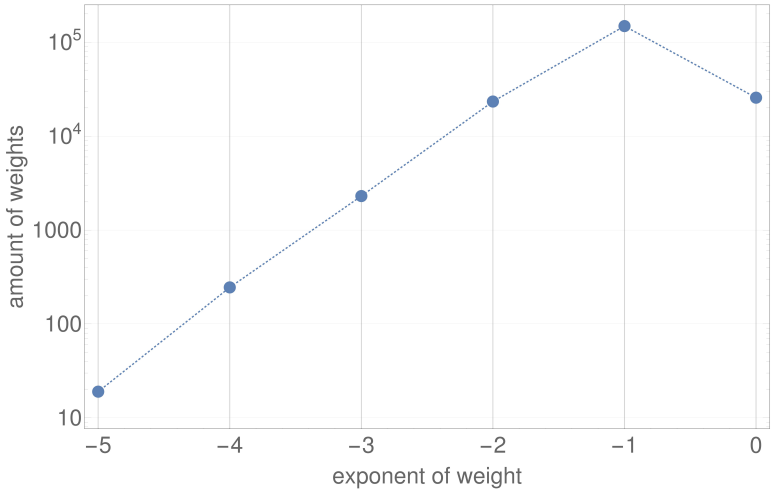


图5：具有 7×100个隐藏层的网络中权重指数的案例数（对数刻度）。该网络的关键参数为：求解器=*L-BFGS*、最大_迭代次数=$5{\cdot}10^{11}$、阿尔法=0.005、学习_速率=逆缩放以及激活函数=*ReLU*。

用作该网络数据库的正电子湮没寿命谱总数为7973，其中6500个用于网络学习输出值（训练集），其余用于检验学习结果（测试集）。表3展示了最终采用的某个网络随机给出的若干结果示例。将训练后的网络给出的结果与LT分析得出的预期值进行了比较。

表3：随机选取的若干计算结果（预测）示例，将 I_2、 τ_2和 τ_3 参数与通过LT计算得到的预期值进行了比较。此处，隐藏_层_大小=7×150，训练集和测试集均为 $\mathcal{S}$=0.985。

示例	I_2 [%]		τ_2 [纳秒]		τ_3 [纳秒]	
	预期	预测值	预期的	预测值	预期的	预测值
1	68.0	68.8	0.27	0.28	1.21	1.20
2	47.2	47.6	0.38	0.39	3.23	3.18
3	65.4	65.5	0.30	0.29	1.35	1.38
4	78.3	78.3	0.23	0.24	1.11	1.12
5	52.4	52.4	0.35	0.34	2.93	2.93
6	60.5	60.4	0.31	0.30	1.25	1.20
7	59.3	58.8	0.29	0.29	1.19	1.22
8	61.0	62.3	0.30	0.31	1.91	1.91
9	70.2	69.5	0.21	0.21	1.06	1.10
10	39.9	38.8	0.23	0.22	1.15	1.13

尽管在本例中，训练集和测试集的 $\mathcal{S}$ 并非我们测试中取得的最高值，但从实际角度来看，使用该网络的结果令人满意，因为预测结果与预期结果之间的偏差处于 LT 本身给出的偏差范围内。

值得一提的是，使用多层感知机回归进行计算前，光谱需要预先处理。主要问题是应在何处截取光谱，以及通过对连续数值取平均来平滑光谱，平滑到何种程度才可以接受。对于第一个问题，通过一系列实验确定：将光谱截取在不同于所述 2k 个通道的界限处后，发现 2k 个通道几乎是最佳选择。 $\langle\mathcal{S}\rangle$在截取范围较短时（例如 1.5k 个通道）会变差，而范围较长时（例如 3k 个通道）则没有显著改善（但由于神经元数量增加，计算结果所需时间更长）。

当学习和测试过程仅限于一种 Δ，并且网络的所有参数保持不变时，预测的准确度会提高。在这种情况下，每条光谱的 t部分中的数值范围要窄得多（只有 δ 会变化）。即使训练集规模缩小，此时训练过程也更有效。为此，我们从仅测量了一种Δ=11.9 ps 的光谱中分出一组。于是，所研究的样品总数缩减为 4116 个，其中 3000 个在经过上述变换后用于训练。在这种情况下得到的得分 $\mathcal{S}$ 远高于将相同网络应用于包含所有可能 Δ的光谱时得到的得分 $\mathcal{S}$。这两种情况的比较见表4（表格最后一行）。为确保不同初始权重（随机赋值）的网络所得结果可靠，此处的得分取 30 次运行的平均值。

表4：不同输入数据格式下 MLPR 验证得分的比较 $\mathcal{S}$。比较“原始”数据（按照第 3 节所述流程对数据进行归一化和调整后取对数）与经过移动平均和压缩平均处理的数据（分别以每 3 个和 5 个独立光谱点为一组）。表格最后一行补充了以某一选定 Δ 所采集数据作为输入的网络结果。

MLPR 与光谱参数			$\langle \mathcal{S}_{tr} \rangle$	$\langle \mathcal{S}_{te} \rangle$
求解器=*L-BFGS* 激活函数=*ReLU* 学习_率=逆缩放 阿尔法=0.01, In=800 隐藏_层_大小=7×150 max_迭代=5×10^9 对 30 次试验取平均	若干个 Δ N_{tr}=6500(∼ 80%) N_{te}=1473	未平滑数据	0.984±0.005	0.963±0.006
		移动平均	0.984±0.005	0.977±0.007
		k=3	0.981±0.005	0.976±0.007
		k=5	0.981±0.006	0.974±0.010
	固定 Δ=11.9 正电子素 N_{tr}=3000 (∼73%) N_{te}=1116	k=5	0.993±0.002	0.989±0.003

验证得分 $\mathcal{S}$ 对光谱平滑处理很敏感，而平滑处理在一定程度上会减少正电子湮没寿命谱所提供的信息。表4中比较了两种情况：分别取光谱中每3个和每5个数据点组成的非重叠窗口，并计算其平均振幅。例如，N=3500 个初始光谱数据点在对每 k=5 个数据点进行平均后，会减少至700个数据点；当 N 除以 k 的余数不为零时，取整数商。$\langle \mathcal{S} \rangle$ 对这两种情况的计算结果表明，从统计上看，两者的结果相同。然而，进一步缩减光谱，即将k=6设为更大的值，会导致 $\mathcal{S}$变差。

表4还显示了在光谱制备过程中应用移动平均时的 $\mathcal{S}$ 参数。此处使用的采样窗口为10。将这一结果与使用未平滑数据进行计算的结果相比，可以发现，应用移动平均确实改善了测试数据集的预测结果。

5 结论

本文表明，易于获取的机器学习工具 MLPRegressor，辅以一些 Python 编程进行数据预处理，可作为解决 PALS 谱反演问题的一种替代方法。所述方法的主要缺点是，需要借助其他软件对训练光谱进行分解，以获得用于训练的输出值。收集好训练集并训练网络后，该算法运行速度非常快，可给出测试光谱的结果。本文采用的训练过程基于 LT 给出的结果，即一种本身会产生一定不确定性的方法。LT 产生的不确定性源于在特定情况下使用数值方法进行拟合计算。另一方面，由于 MLPR 预测基于大量光谱的信息，这种方法似乎对给定光谱的具体形状不太敏感，因此预测参数的准确性可能更高。此外，所述方法似乎比文献中提及的方法更快，因为训练后的网络只需进行简单的矩阵和向量变换即可完成计算，计算开销较低，也较不容易受到数值问题的影响。

尽管此处提出的模型与 [14] 中描述的模型相似（并在 [15]中再次提及），但两者存在表5所示的显著差异。我们的实验数据由功能特性各异的光谱仪采集，尤其是时间分辨率不同。即使对于同一台光谱仪，由于实验条件（特别是温度）发生变化，也应定期重新校准这一参数。在 [14] 中提出的算法

表5：本研究中应用的 MLPR 建模（$skLearn$）与此前发表的三组分光谱分析（见 [14] 和 [15]）的关键参数和结果比较。

	Pázsit [14]	一个 [15]	**skLearn**
训练光谱类型	模拟的	模拟	真实（烷烃）
训练光谱数量	575	920	7973
测试光谱类型	模拟	模拟、硅	烷烃
测试光谱数量	50	100（30）	1473
网络类型	单层感知器	单层感知器	多层感知器
道宽 [皮秒]	23.2	24.5	**一些(11.2-19.5)**
多道分析器/采集道数	1500/1500	1024/1024	8192/3500
光谱中的计数近似值	10M	10M	∼400千
求解器	反向误差传播	反向误差传播	**some**
隐藏层数量	1	1	some
I_2, τ_3 平均误差 [%] 在测试的**模拟**光谱上	7.3, 1.0	1.07-3.52, 0.55-1.21	-, -
I_2, τ_3 平均误差 [%] 在测试的**真实**光谱上	-, -	-, -	1.03, 1.70

假设所有光谱使用相同的分辨曲线。在我们的数据准备流程中，会针对每种情况对分辨曲线的参数进行插值。基于此，计算 δ 参数，并确定连续道的时间偏移值。此处假设分辨曲线为单高斯分辨率曲线，但该算法也可以扩展到更复杂的情况，例如分布曲线由多个高斯函数之和构成。正如 [14], 中所述，在这种情况下，若能识别分布函数，便可与 MELT [23]兼容。要实现这种扩展，需要通过应用另一个预先运行的神经网络来扩展计算，该网络可返回相应情况下的分辨曲线参数。这个问题已通过应用霍普菲尔德神经网络 [16]得到解决。考虑到我们收集的光谱数据，我们使用 LT（偶尔也使用 MELT）进行了检查，结果表明，对于我们的光谱，仪器分辨率曲线为单高斯曲线。因此，这些数据无法用于测试这种扩展模型。

此外，光谱仪的多道分析器模块的每道时间常数可能不同 Δ。因此，用作训练数据集的光谱可能是在不同道宽下采集的。考虑到 [14] 中针对固定 Δ 提出的方法，对于使用其他 Δ 采集的光谱，训练结果用处不大。与此相反，我们已证明，改进的算法可以应用于针对不同 Δ 采集的数据。来自多台光谱仪的数据有助于构建大型训练数据集，从而能够针对任何正电子湮没寿命谱求解反演问题，因此可成为一种适用于不同实验室的通用工具。尽管此处使用的 Δ 数量较少，结果的准确度仍相当不错。若要使用这一工具确定实际谱参数，应通过添加在更广泛的 Δ 范围内测得的光谱来扩充训练过程。

为了提高泛化能力，原则上可以将针对其他化合物采集的光谱添加到训练数据集中。为保持一致，训练数据库只需在谱分解中采用相同数量的组分（此处为三个）。然而，在实践中，不同化合物的光谱可能存在某些不兼容之处，因为将光谱分解为少数几个指数过程，或许始终只是对真实情况的一种简化；真实情况还应考虑自由体积尺寸和形状分布，以及正电子素形成的其他细节。

尽管本文提出的方法仅限于烷烃分析，但该算法也可用于计算其他类型样品的正电子湮没寿命谱参数。

参考文献

[1] D. Manara, A. Seibert 等。正电子湮没谱学。载于 M.H.A. Piro 编，《核燃料化学进展》，第 2.3.5 章，第 126–131 页。爱思唯尔有限公司、伍德海德出版社，2020年。

[2] W. Greiner 和 J. Reinhardt。《场量子化》。施普林格，2013年。

[3] T. Goworek。正电子素可作为晶体、聚合物和多孔介质中小自由体积的探针。Ann. Univ. Mariae Curie Sklodowska, sectio AA – Chemia，LXIX:1–110，2014年。[4] Y.C. Jean、J.D. Van Horn、W-S Hung 和 K-R Lee。正电子湮没谱学在聚合物中的前景。Macromolecules，46:7133–7145，2013年。[5] J. Kansy。用于分析正电子湮没寿命谱的微型计算机程序。Nucl. Instr. Meth. A，374:235–244，1996年。[6] Jens V. Olsen、Peter Kirkegaard、Niels Jørgen Pedersen 和 Morten Mostgaard Eldrup。Palsfit：一种用于评估正电子寿命谱的新程序。Physica Status Solidi (C) Current Topics in Solid State Physics，4(10):4004–4006，2007年。[7] K. Wada 和 T. Hyodo。一种用于正电子湮没寿命谱学孔径估算的简单无形状模型。J. Phys. Conf. Ser.，443:012003，2013年。[8] J. Jegal、D. Jeong、E.S. Seo 等。基于卷积神经网络的正电子素湮没定位重建。Sci Rep，12:8531，2022年。[9] J.L. Herraiz、A. Bembibre 和 A. López-Montes。基于深度学习的PET图像正电子射程校正。Applied Sciences，11(1)，2021年。[10] M. Wędrowski。基于人工神经网络的正电子发射断层成像位置估计。博士论文，Interuniversity Institute for High Energies，Vrije Universiteit Brussel，比利时，2010年。[11] W.J. Whiteley。《正电子发射断层成像中的深度学习图像深度学习重建重建》。博士论文，田纳西大学，诺克斯维尔，美国，2020年。[12] D. Petschke 和 T.E.M. Staab。一种用于正电子湮没寿命谱学（pals）中形状敏感探测器脉冲甄别的监督式机器学习方法，采用朴素高斯贝叶斯分类。Nucl. Instrum. Methods. Phys. Res. B，947:162742，2019年。[13] N.H.T. Lemes、J.P. Braga 和 J.C. Belchior。遗传算法在反演正电子寿命光谱中的应用。Chem. Phys. Lett.，412(4):353–358，2005年。[14] I. Pázsit、R. Chakarova、P Lindén 和 F. Maurer。利用神经网络展开正电子寿命光谱。Appl. Surf. Sci.，149:97–102，1999年8月。[15] R. An、J. Zhang、W. Kong 和 B-J. Ye。人工神经网络在正电子寿命光谱反演中的应用。Chinese Physics B，21(11):117803，2012年11月。[16] V.C. Viterbo、J.P. Braga、A.P. Braga 和 M.B. de Almeida。利用神经网络反演模拟正电子湮没寿命谱。J. Chem. Inf. Comput. Sci.，41:309–313，2001年。[17] G. Rebala、A. Ravi 和 S. Churiwala。《机器学习导论》。Springer Nature，2019年。[18] J. Heaton。《神经网络数学导论》。Heaton Res.，2012年。[19] F. Pedregosa、G. Varoquaux、A. Gramfort、V. Michel、B. Thirion、O. Grisel、M. Blondel、P. Prettenhofer、R. Weiss、V. Dubourg 等。Scikit-learn：Python中的机器学习。J. Mach. Learn. Res.，12:2825–2830，2011年。[20]Broyden–fletcher–goldfarb–shanno算法。https://en.wikipedia.org/wiki/Broyden-Fletcher-Goldfarb-Shanno_算法。访问日期：2022年11月20日。[21]sklearn.neural_network.mlpregressor。https://scikit-learn.org/stable/modules/generated/ sklearn.neural_network.MLPRegressor.html。[22] T. Goworek、M. Pietrow、R. Zaleski 和 B. Zgardzińska。长链偶数正构烷烃高温相中的正电子素。Chem. Phys.，355:123–129，2009年。[23] A. Shukla、M. Peter 和 L. Hoffmann。利用量化最大熵和一般线性滤波器分析正电子寿命光谱。Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment，335(1):310–317，1993年。