arXiv:2302.02403v2 [cs.CE] 6 Jul 2023

# 神经网络邂逅超弹性：物理约束实施指南

Lennart Linden[1], Dominik K. Klein[2], Karl A. Kalina[1], Jörg Brummund[1],OliverWeeger[2], 和 Markus Kästner[1,3*]

[1]德累斯顿工业大学固体力学研究所，德国德累斯顿 *01062*[2]达姆施塔特工业大学信息物理仿真组与计算工程研究生院、机械工程系与计算工程中心，德国达姆施塔特 *64293*[3]德累斯顿工业大学德累斯顿计算材料科学中心 *(DCMS)*，德国德累斯顿 *01062*[*]通讯作者，电子邮箱：*markus.kaestner@tu-dresden.de*

2023年5月3日

**摘要**

本文提出了一种基于神经网络的超弹性本构模型，该模型通过构造满足所有常见的本构条件，尤其适用于可压缩材料行为。通过采用不同的不变量集作为输入，将超弹性势构建为凸神经网络，从而满足应力张量对称性、客观性、材料对称性、多凸性和热力学一致性。此外，通过使用解析增长项以及归一化项，确保模型具有物理上合理的应力行为；这些归一化项使未变形状态下应力为零且能量为零。特别是，针对各向同性和横观各向同性材料行为，分别构造了多凸、基于不变量的应力归一化项。通过精确满足所有这些条件，所提出的物理增强模型将坚实的力学基础与神经网络所具有的非凡灵活性结合起来，从而使过去数十年发展起来的超弹性理论与最新的机器学习技术相协调。此外，通过对可容许变形状态空间进行采样，数值考察了基于神经网络的超弹性势的非负性；据作者所知，对于所考虑的非线性可压缩模型，这是唯一可行的方法。对于各向同性神经网络模型，通过解析考量缩小了所需的采样空间。此外，本文还证明了可压缩Neo-Hooke势的非负性。通过使用解析势生成的数据对模型进行标定，展示了该模型的适用性，随后将其应用于有限元仿真。此外，本文还展示了模型对含噪数据的适配，并将其外推能力与物理基础较弱的模型进行了比较。在所有数值示例中，所提出的物理增强神经网络均取得了出色且符合物理规律的预测结果。

**关键词：** 超弹性、物理增强神经网络、归一化、各向异性、本构建模、有限元模拟

发表于 *Journal of the Mechanics and Physics of Solids* 的手稿接收版本。接收日期：2023年6月14日。数字对象标识符：10.1016/j.jmps.2023.105363。许可协议：知识共享署名—非商业性使用—禁止演绎 4.0

# 1 引言

过去几十年间，超弹性背后的力学原理得到了广泛讨论，但长期以来，要同时满足所有这些原理，一度被认为是“材料行为理论中的主要未解决问题”（Truesdell 和 Noll [58]）。例如，尽管多凸性 [5, 6] 和客观性条件都具有坚实的力学依据，但客观应变度量却很容易违反多凸性条件 [26]。这并不意味着不同的本构条件彼此矛盾——而是说明要同时满足所有这些条件面临着巨大挑战。随着模型需要满足的限制条件不断增加，实现这一目标的难度也显著提升；人们需要采用复杂的方法，才能构建出同时满足所有相关条件的解析模型 [9, 49]。此外，此类模型的标定也并非易事，需要大量知识 [47]。

为了克服经典本构模型构建过程耗时的问题，并改进大多数此类解析模型所伴随的受限函数关系，数据驱动的力学方法 [25] 或高斯过程回归[12, 14] 等现代机器学习方法，以及神经网络（神经网络） [1, 28] ，都代表了颇具前景的替代方案。20世纪90年代初，Ghaboussi等人首次提出了将神经网络应用于本构建模的想法 [16]。然而，在这一早期阶段，主要采用的是纯黑箱方法，即不考虑任何物理原理的网络，因此只能较好地复现训练数据集（此处由应力-应变对组成），而外推能力较差。为弥补这一缺陷，神经网络本构建模以及整个科学机器学习领域近年来出现了一个相当新的趋势 [46], : 以强形式纳入基本的底层物理学，例如采用经过调整的网络架构；或以弱形式纳入，例如修改训练过程中的损失项 [37]。这类方法被称为物理信息驱动 [24], 力学信息驱动 [4],物理增强 [27], 物理约束 [21], 或基于热力学 [40], ，能够提升外推能力，并实现对稀疏训练数据的利用 [24, 29], ；当需要根据实验数据拟合本构模型时，这一点尤为重要。

下面简要概述上述应用于有限应变超弹性 建模的、基于神经网络的方法。对于各向同性材料，受解析模型启发，文献[32, 51] 提出使用以前馈神经网络（FFNN）近似弹性势，并将三个变形型不变量作为输入，从而满足若干本构条件，例如热力学一致性、客观性或材料对称性。然而，与应用于各向异性问题的方法 [30, 34, 48]类似，在 [32, 51]中进行训练时，需要直接使用弹性势。与此同时，将不变量 作为输入、将超弹性势作为输出的神经网络也已成为一种相当成熟的方法 [13, 21, 26, 27, 33, 35, 36, 55]，因此该方法在先验上也满足热力学一致性。在这种方法中，训练过程更为精细，可以通过应力与应变的元组直接标定网络；也就是说，将能量对变形的导数用于损失项。这种技术也称为 Sobolev训练 [59, 61]。另一种方法是，为了在事后确保热力学一致性，可以使用一个经过预训练、用于预测应力系数的网络来构造伪势 [22]。文献 [66]提出了一种与特定模型（即所谓的微球方法）相结合的神经网络方法。其他方法则根据变形梯度的特征值构建神经网络势 [53], ，从而推广了 Ogden型模型 [45], ；或者根据应变张量的分量构建神经网络势 [4, 10, 26, 61]。

除了上述要求，即热力学一致性、客观性和材料对称性之外，还存在其他物理条件，例如椭圆性，它可确保材料稳定性 [64]。然而，由于椭圆性难以验证和保证，因此在构建本构模型时，优先采用应变能势的多凸性概念[5, 6],；多凸性蕴含椭圆性，并且在数学上与弹性问题解的存在性和稳定性相关 [43]。构建多凸神经网络 [8, 26, 36, 54, 55], 有多种方法，其中纳入这一条件最著名的技术是使用输入凸神经网络（*ICNN*），该网络最初由 Amos等人提出 [3]。

其他方法则考虑应变能势关于右柯西-格林变形张量的凸性 $\boldsymbol{C}$。除了多凸性之外，这或许也是应变能势值得具备的特性[31], , 但仅关于 $\boldsymbol{C}$ 的凸性不足以确保材料稳定性 [63]。与多凸神经网络类似，关于 $\boldsymbol{C}$ 凸的应变能势可以用输入凸神经网络表示 [4]。在 [59], 中，针对一个特定数据集考察了所提出势函数关于 $\boldsymbol{C}$ 的凸性。对于所考察的情形，所提出的势函数确实关于 $\boldsymbol{C}$凸；但这一条件并非通过构造得到保证，因此在模型的其他应用中可能会被违反。为满足增长条件，可以采用特殊的网络架构 [21], , 而使用解析增长项则更为普遍 [26, 27]。最后，虽然有多项研究引入了用于确保满足归一化条件的修正项，但这些修正项要么不具备多凸性 [4, 10, 57] ，要么仅适用于近不可压缩材料行为 [36]。例如， [36] 中提出的模型包含如下形式的项：$(I_1 - 3)^2$, , 其中 $I_1$ 表示右柯西-格林变形张量的第一不变量。然而，为了保持 $I_1$的多凸性，作用于该项的函数通常必须是凸且单调不减的，参见 [26,附注 A.10]。对于 [36], 中使用的二次函数，只有当 $I_1$ 不小于 3 时，这一条件才成立；这仅在 $\det \boldsymbol{F} = 1$时满足，参见 [49, 推论 A.11], , 其中 $\boldsymbol{F}$ 表示变形梯度。因此，对于 $\det \boldsymbol{F} \neq 1$，势函数关于 [36] 的多凸性并未通过构造得到保证。总的来说，据我们所知，迄今文献中的模型只能精确满足可压缩超弹性的一部分本构条件，其余条件则只是近似满足，也就是说，它们通过损失函数中的惩罚项予以考虑 [56, 62]。更具体地说，虽然 [26] 以精确方式满足多凸性条件，但归一化条件仅通过标定数据进行学习，从而得到近似满足。另一方面，虽然 [57] 使用应力修正来精确满足归一化条件，但这一应力修正项包含右柯西-格林变形张量的非对角分量，因此通常不具备多凸性，且会违反材料对称性。

总而言之，对于基于神经网络的本构模型，如何同时精确满足可压缩超弹性的所有常见本构条件，至今仍是一项尚未解决的挑战。[1] 特别是，据作者所知，迄今尚未提出一种用于描述可压缩材料行为的多凸应力归一化方法，既不违反角动量平衡、客观性，也不违反材料对称性条件。本文系统地推导出一种此类方法，通过构建网络架构，同时确保热力学一致性、应力张量的对称性、客观性、材料对称性、多凸性、增长条件以及能量和应力归一化。对于所提出的新型基于神经网络的超弹性模型系列——它们以精确方式满足上述所有条件——我们建议将其命名为物理增强神经网络（*PANN*）。对于需要高度灵活且同时符合物理规律的本构模型的领域，例如微结构材料的模拟，所提出的框架将具有重要价值[17, 21]。为此，PANN方法通过扩展上述模型 [26]构建而成。通过构造多凸的、基于不变量的应力归一化项，补上了可压缩情形下多凸性与应力归一化之间缺失的一环。正如注3.2所指出的，这些条件尤其难以结合。随后，本文对整体模型进行了详细的解析和数值研究，包括对不同模型方法的广泛考察，范围从不包含任何力学条件的朴素神经网络模型，到包含所有力学条件的PANN模型。本文全面考察了不同模型的插值和外推表现，并进行了统计评估。此外，文中逐步清晰地说明了如何将所考虑的所有物理条件纳入PANN方法；该方法尤其适用于可压缩各向异性材料行为。我们将该框架应用于各向同性和横观各向同性情形，并考察了若干具有代表性的示例，包括多轴应力-应变状态和含噪数据。

[1] 这仅适用于可压缩弹性材料的情形。Linka 和 Kuhl [36] 针对不可压缩情形提出了相应方案。然而，除了局限于不可压缩材料之外，与本文介绍的方法相比，该方案在其他一些方面的通用性也较低。

训练所用的数据基由解析势生成。对模型进行标定后，将其用于有限元计算，以展示其适用性和总体准确性。最后，本文给出了可压缩Neo-Hooke势非负性的证明；据作者所知，此前尚未有人完成这一证明。

本文的结构如下：第 2 节讨论超弹性力学基础，并在第 3 节将其应用于所提出的神经网络模型。第 4 节给出数值示例。最后，第 5 节通过简短讨论以物理学增强神经网络的重要性作结。除了一些辅助性附录外，附录 C 还给出了可压缩Neo-Hooke势非负性的证明。

**符号** 在全文中，张量空间

$$\mathcal{L}_n := \underbrace{\mathbb{R}^3 \otimes \cdots \otimes \mathbb{R}^3}_{n\text{-times}} \ \forall n \in \mathbb{N} \tag{1}$$

均用于表示张量，但秩为零的张量除外。在公式 (1) 中，$\mathbb{R}^3$、$\mathbb{N}$ 和 $\otimes$ 分别表示欧几里得向量空间、不包含零的自然数集合以及并矢积。下文中，秩为一和秩为二的张量分别用粗斜体符号表示，即 $\boldsymbol{a} \in \mathrm{L}_1$ 或 $\boldsymbol{B}$、$\boldsymbol{C} \in \mathrm{L}_2$。二阶张量 $\boldsymbol{B}$ 的转置和逆分别用 $\boldsymbol{B}^T$ 和 $\boldsymbol{B}^{-1}$ 表示。此外，迹、行列式和余子式分别记为 tr $\boldsymbol{B}$、det $\boldsymbol{B}$ 和 cof $\boldsymbol{B} := \det(\boldsymbol{B})\, \boldsymbol{B}^{-T}$。具有正行列式的可逆二阶张量集合记为 $\mathcal{GL}^+(3) := \{\boldsymbol{A} \in \mathrm{L}_2 \,|\, \det \boldsymbol{A} > 0\}$，而 $\mathbb{R}^3$ 中的正交群和特殊正交群分别记为 $\mathcal{O}(3) := \left\{\boldsymbol{A} \in \mathrm{L}_2 \,|\, \boldsymbol{A}^T \cdot \boldsymbol{A} = \boldsymbol{1}\right\}$ 和 $\mathcal{SO}(3) := \left\{\boldsymbol{A} \in \mathrm{L}_2 \,|\, \boldsymbol{A}^T \cdot \boldsymbol{A} = \boldsymbol{1},\ \det \boldsymbol{A} = 1\right\}$。这里，$\boldsymbol{1} \in \mathrm{L}_2$ 表示二阶单位张量。对称二阶张量空间记为 $\mathcal{Sym} := \left\{\boldsymbol{A} \in \mathrm{L}_2 \,|\, \boldsymbol{A} = \boldsymbol{A}^T\right\}$。此外，两个张量的单缩并和双缩并分别表示为 $\boldsymbol{B} \cdot \boldsymbol{C} = B_{kq} C_{ql} \boldsymbol{e}_k \otimes \boldsymbol{e}_l$ 和 $\boldsymbol{B} : \boldsymbol{C} = B_{kl} C_{lk}$。其中，$\boldsymbol{e}_k \in \mathrm{L}_1$ 表示笛卡尔基向量，并采用爱因斯坦求和约定。$\nabla_{\boldsymbol{X}}$ 是关于参考构形 $\mathrm{B}_0$ 的纳布拉算子。为便于阅读，本文通常省略函数的自变量。然而，为了展示函数之间的依赖关系，除书写导数时外，能量函数均给出其自变量。此外，我们令函数的符号与函数值本身的符号相同。

# 2 超弹性基础

## 2.1 运动学、应力度量和平衡方程

考虑一个固体的运动，在时刻 $t_0 \in \mathbb{R}$，其参考构形为 $\mathrm{B}_0 \subset \mathbb{R}^3$。变形后，该物体在时刻 $t \in \mathrm{T} := \{\tau \in \mathbb{R} \,|\, \tau \geq t_0\}$ 的当前构形为 $\mathrm{B} \subset \mathbb{R}^3$。物体的运动由双射映射 $\boldsymbol{\varphi} : \mathrm{B}_0 \times \mathrm{T} \to \mathrm{B}$ 定义，该映射将物质粒子 $\boldsymbol{X} \in \mathrm{B}_0$ 与 $\boldsymbol{x} = \boldsymbol{\varphi}(\boldsymbol{X}, t) \in \mathrm{B}$ 联系起来。与 $\boldsymbol{\varphi}$ 相对应，变形梯度 $\boldsymbol{F} \in \mathcal{GL}^+(3)$ 及其行列式定义为 $\boldsymbol{F} := (\nabla_{\boldsymbol{X}} \boldsymbol{\varphi})^T$ 和 $J := \det \boldsymbol{F} \in \mathbb{R}_{>0}$。一种不受刚体运动影响的变形度量由正定的右柯西-格林变形张量 $\boldsymbol{C} := \boldsymbol{F}^T \cdot \boldsymbol{F} \in \mathcal{Sym}$ 给出。

根据非线性连续介质力学的概念，我们引入对称柯西应力$\boldsymbol{\sigma} \in \mathcal{Sym}$ 以及第一类和第二类皮奥拉-基尔霍夫应力张量 $\boldsymbol{P} \in \mathrm{L}_2$ 和 $\boldsymbol{T} \in \mathcal{Sym}$。后两者分别通过拉回运算 $\boldsymbol{P} := J\boldsymbol{\sigma} \cdot \boldsymbol{F}^{-T}$ 和 $\boldsymbol{T} := \boldsymbol{F}^{-1} \cdot \boldsymbol{P}$得到。

为补全材料无关方程组，简要列出相关平衡定律。考虑由 $\varrho_0 = J\varrho$ 给出的质量守恒，其中 $\varrho_0$ 和 $\varrho$ 分别表示参考构形和当前构形的质量密度，则相对于 $\mathrm{B}_0$ 的线动量平衡为

$$\nabla_{\boldsymbol{X}} \cdot \boldsymbol{P}^T + \varrho_0 \boldsymbol{f} = \boldsymbol{0}\,. \tag{2}$$

在上述方程中，忽略了惯性项，$\boldsymbol{f} \in \mathrm{L}_1$ 表示单位质量力密度。此外，考虑质量平衡和线动量平衡式 (2)，可得角动量平衡

$$\boldsymbol{P} \cdot \boldsymbol{F}^T = \boldsymbol{F} \cdot \boldsymbol{P}^T \,. \tag{3}$$

因此，利用上述应力变换可知，角动量守恒要求应力张量 $\boldsymbol{\sigma}$ and $\boldsymbol{T}$对称。所引入的力学平衡方程还需补充边界条件 $\boldsymbol{u} = \hat{\boldsymbol{u}}$ on $\partial \mathrm{B}_0^{\boldsymbol{u}}$and $\boldsymbol{p} = \hat{\boldsymbol{p}}$ on $\partial \mathrm{B}_0^{\boldsymbol{p}}$，其中 $\partial \mathrm{B}_0^{\boldsymbol{u}}$and $\partial \mathrm{B}_0^{\boldsymbol{p}}$表示本质边界和自然边界，并满足 $\partial \mathrm{B}_0^{\boldsymbol{u}} \cup \partial \mathrm{B}_0^{\boldsymbol{p}} = \partial \mathrm{B}_0$and $\partial \mathrm{B}_0^{\boldsymbol{u}} \cap \partial \mathrm{B}_0^{\boldsymbol{p}} = \varnothing$。最后，结合前面引入的平衡方程以及能量平衡和熵，可根据热力学第二定律推出克劳修斯–杜亥姆不等式。若忽略热效应，该关系可表示为

$$-\dot{W} + \boldsymbol{P} : \dot{\boldsymbol{F}}^T \geq 0 \,, \tag{4}$$

¤ 其中，$W$ 表示相对于 $\mathrm{d}V_0 \subset \mathrm{B}_0$ 的亥姆霍兹自由能密度，$(\cdot)$表示物质时间导数。

## 2.2 超弹性的总体要求

弹性力学中的本构建模旨在建立材料点处应变与其所引发的应力之间的关系。在超弹性中，应变与应力之间的映射并非直接定义，而是引入一种势函数

$$\psi : \mathcal{GL}^+(3) \to \mathbb{R}, \quad \boldsymbol{F} \mapsto \psi(\boldsymbol{F}) \tag{5}$$

，它对应于物体中储存的应变能密度，且等于亥姆霍兹¤自由能密度 $W$。因此，为满足任意 $\boldsymbol{F}$ 下的不等式 (4)，相应的第一类 Piola-Kirchhoff 应力为

$$\boldsymbol{P} = \frac{\partial \psi}{\partial \boldsymbol{F}} \,. \tag{6}$$

根据此定义，应力张量是一个梯度场，这意味着能量守恒和路径无关性，因此其构造天然具备**热力学一致性** [21, 33, 36]。[2]

超弹性势还需满足进一步的数学和物理考量，以下将对此作简要讨论，并在图 1 中加以说明。有关超弹性的更详细介绍，请参阅 [18, 19, 47]。

首先，必须构造 $\psi(\boldsymbol{F})$，以确保其与角动量平衡相容。利用式（3）和（6），这一要求可表示为

$$\frac{\partial \psi}{\partial \boldsymbol{F}} \cdot \boldsymbol{F}^T = \boldsymbol{F} \cdot \frac{\partial \psi}{\partial \boldsymbol{F}^T} \,. \tag{7}$$

利用第 2.1 节介绍的应力变换，这要求**应力张量的对称性** $\boldsymbol{\sigma}$ 和 $\boldsymbol{T}$，例如 $\boldsymbol{T} = \boldsymbol{F}^{-1} \cdot \frac{\partial \psi}{\partial \boldsymbol{F}} \mathrel{!}= \frac{\partial \psi}{\partial \boldsymbol{F}^T} \cdot \boldsymbol{F}^{-T} = \boldsymbol{T}^T$。

[2]请注意，本构方程必须始终符合物理规律，即不得与所引入的平衡方程相矛盾。与热力学第二定律相容尤为重要，这里该定律以克劳修斯–杜亥姆不等式（4）的形式给出。因此，这里特别强调热力学一致性。

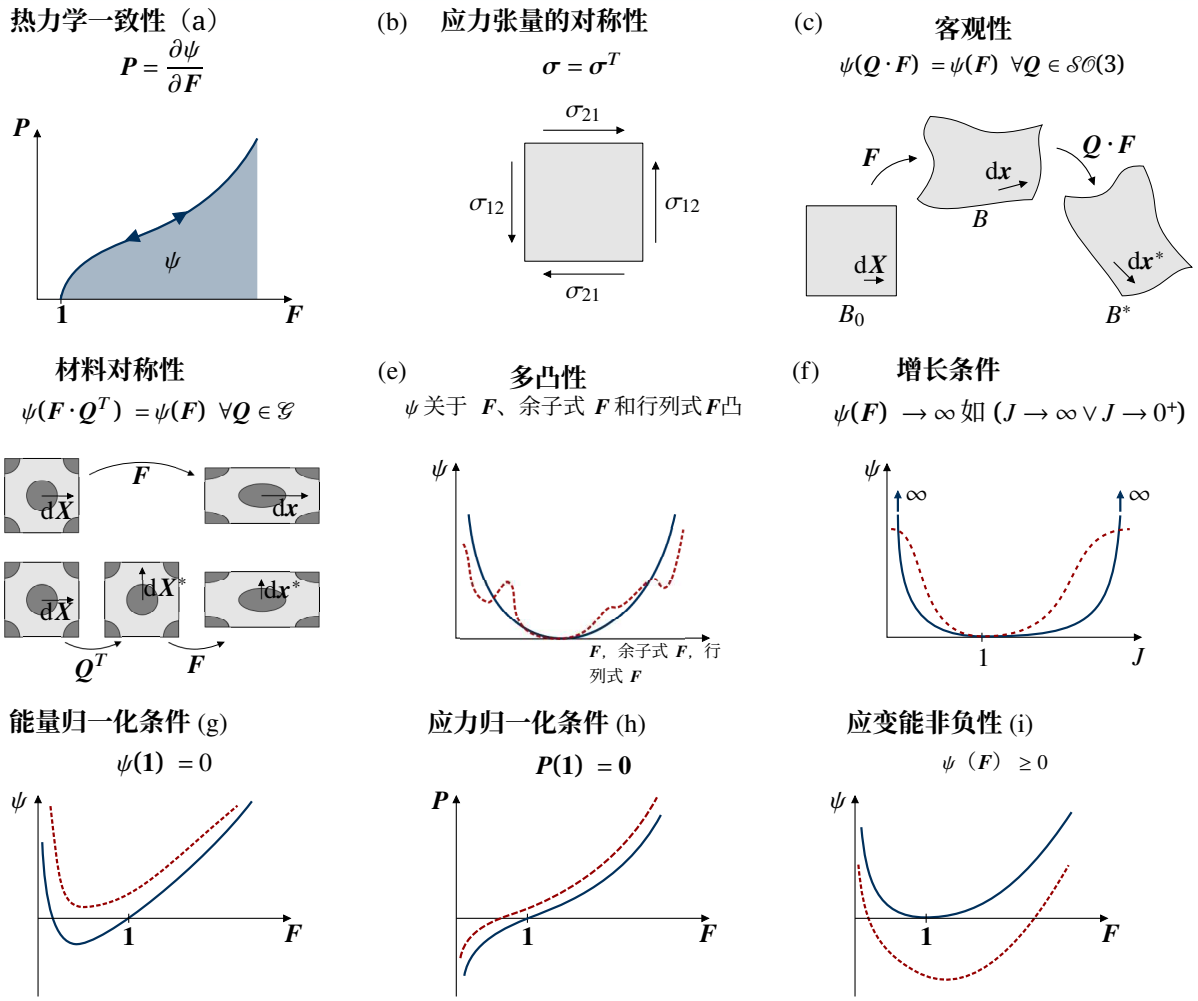


图1：弹性势和应力常见条件的示意图：(a) 热力学一致性，(b) 柯西应力的对称性，(c) 客观性，(d) 材料对称性，(e) 多凸性，(f) 体积增长条件，(g) 和 (h) 能量和应力的归一化条件，以及 (i) 能量非负性。在 (e)–(i) 中，蓝色实线表示满足相应条件的模型，红色虚线则表示违反相应条件的模型。在 (c) 和 (d) 中，d$\boldsymbol{X}$、d$\boldsymbol{X}^*$、d$\boldsymbol{x}$ 和 d$\boldsymbol{x}^*$ 表示不同变形状态下的材料线元，其中 d$\boldsymbol{X}^*$ 和 d$\boldsymbol{x}^*$ 分别经过 d$\boldsymbol{X}$ 和 d$\boldsymbol{x}$ 旋转变换。

其次，本构模型应不受观察者选择的影响，这称为**客观性**。在超弹性理论中，这一性质被形式化为

$$\psi(\boldsymbol{Q}\cdot\boldsymbol{F}) = \psi(\boldsymbol{F}) \quad \forall \boldsymbol{F} \in \mathcal{GL}^+(3), \boldsymbol{Q} \in \mathcal{SO}(3). \tag{8}$$

本构方程还应反映材料固有的各向（异）性，这称为**材料对称性**，数学表达式为

$$\psi(\boldsymbol{F}\cdot\boldsymbol{Q}^T) = \psi(\boldsymbol{F}) \quad \forall \boldsymbol{F} \in \mathcal{GL}^+(3), \boldsymbol{Q} \in \mathcal{G} \subseteq \mathcal{O}(3), \tag{9}$$

其中，$\mathcal{G}$ 表示所考虑材料的对称群。

此外，我们还考虑具有如下表示形式的多凸超弹性势 [5, 6, 9, 49]，

$$\psi(\boldsymbol{F}) = \mathcal{P}(\boldsymbol{F}, \operatorname{cof}\boldsymbol{F}, \det\boldsymbol{F}), \tag{10}$$

其中，$\mathrm{P}(\boldsymbol{F},\ \mathrm{cof}\ \boldsymbol{F},\ \det \boldsymbol{F})$ 关于其参数是凸的，$\mathrm{cof}\ \boldsymbol{F}=\det(\boldsymbol{F})\ \boldsymbol{F}^{-T}$表示变形梯度的余子式。多凸性源于一个相当理论化的背景——然而，它意味着椭圆性[64], ，这在实践中十分重要，而且与多凸性相比，将其纳入模型公式要困难得多。椭圆性（或秩一凸性）条件 [43, 64]

$$(\boldsymbol{a}\otimes\boldsymbol{b}):\frac{\partial^2\psi}{\partial\boldsymbol{F}\partial\boldsymbol{F}}:(\boldsymbol{a}\otimes\boldsymbol{b})\geq 0\qquad \forall\boldsymbol{a},\boldsymbol{b}\in\mathcal{L}_1\,, \tag{11}$$

确保模型的材料稳定性，从而使其在数值应用中表现良好。请注意，为确保多凸性，构建超弹性势的所有步骤都必须保持多凸性，尤其必须使用多凸不变量 [41],参见第 2.3 节。

此外，还可以考虑多种强制性条件，其中最常见的是**体积增长条件**

$$\psi(\boldsymbol{F})\to\infty\quad\text{as}\quad\left(J\to 0^{+}\quad\vee\quad J\to\infty\right), \tag{12}$$

这是为了考虑到这样一种事实：材料体不能被压缩至零体积，也不能膨胀至无限体积 [19]。特别是，大幅体积压缩的情况很重要，因为它可能出现在工程应用的相关范围内。

最后，式（5）给出的函数关系还需进一步满足物理上合理的行为要求。在未变形构形中，即 $\boldsymbol{F}=\mathbf{1}$，应满足**归一化**条件

$$\psi(\boldsymbol{F}=\mathbf{1})\overset{!}{=}0\quad\text{and}\quad\boldsymbol{P}(\boldsymbol{F}=\mathbf{1})\overset{!}{=}\mathbf{0} \tag{13}$$

能量和应力均应满足上述条件。除归一化外，只要发生变形，自由能就应增加。因此，除式（13）外[1], ，还要求**应变能非负**，即$\psi(\boldsymbol{F})\geq 0$。

将势函数表示为由右柯西-格林变形张量 $\boldsymbol{C}$ 和一组结构张量导出的不变量的函数，并将其概括为集合 $\mathrm{S}^{\square}$。该集合反映了所研究材料体的材料对称性 [18, 19], ，即：

$$\psi:\mathbb{R}^m\to\mathbb{R},\quad \boldsymbol{\mathcal{I}}\mapsto\psi(\boldsymbol{\mathcal{I}})\,, \tag{14}$$

同时满足**客观性** 和**材料对称性** 条件 [47]。[3]其中，$\mathrm{I}:=(I_1,\ldots,I_m)\in\mathbb{R}^m$表示一个 $m$元组，其中包含对称群的一组完备且不可约的不变量 $I_\alpha(\boldsymbol{C},\mathrm{S}^{\square})$。这里，完备是指任意不变量均可表示为 I 中各元素的函数，例如 $\psi(\mathrm{I})$。不可约是指集合 I 中的一个不变量 $I_\alpha$无法用 I 中的其他元素表示，参见 [9]。请注意，必须选择这些不变量，使其不违反**多凸性** 条件。[4] 相应的第二类皮奥拉-基尔霍夫应力张量为

$$\boldsymbol{T}=2\frac{\partial\psi}{\partial\boldsymbol{C}}=2\sum_{\alpha=1}^{m}\frac{\partial\psi}{\partial I_\alpha}\frac{\partial I_\alpha}{\partial\boldsymbol{C}}=\boldsymbol{F}^{-1}\cdot\boldsymbol{P}\,. \tag{15}$$

最后，将超弹性势表示为右柯西-格林张量不变量的函数，意味着**柯西应力张量 具有对称性**，从而确保角动量守恒 [19]。

[3]请注意，根据自变量的不同，$\psi(\boldsymbol{F})$和$\psi(\mathrm{I})$是不同的函数。为减少符号数量，本文采用这种非正式记法。[4]需要注意的是，对于某些对称群，无法找到不违背多凸性的完备集[27]。在这种情况下，必须决定模型的哪项性质更为重要。此外，对于具有任意微观结构的材料（例如3D打印复合材料），可能不存在能够描述其各向异性的完备不变量集。

### 2.3 特殊材料对称类别和特定模型

本文以各向同性（$\square := \circledcirc$）和横观各向同性（$\square := \mathrm{k}$）材料行为为例，探讨特定的材料对称群。在各向同性情形下，即材料响应与方向无关时，$\mathcal{G} = \mathcal{O}(3)$给出了一组完备且不可约的不变量，由以下三个不变量组成

$$I_1 := \operatorname{tr}\boldsymbol{C}, \qquad I_2 := \operatorname{tr}(\operatorname{cof}\boldsymbol{C}), \qquad I_3 := \det\boldsymbol{C}\,. \tag{16}$$

在这种情况下，无需结构张量，因此 $\mathsf{S}^{\circledcirc} = \varnothing$ 成立。对于横观各向同性材料，完整的函数基由三个各向同性不变量 $I_1, I_2, I_3$ 以及

$$I_4 := \operatorname{tr}(\boldsymbol{C}\cdot\boldsymbol{G}), \qquad I_5 := \operatorname{tr}(\operatorname{cof}(\boldsymbol{C})\cdot\boldsymbol{G})\,, \tag{17}$$

其中，$\boldsymbol{G}$ 表示二阶横观各向同性结构张量 [9, 50]。当优选方向平行于 $X_1$方向时，结构张量的分量为

$$(G_{KL}) := \begin{pmatrix} \beta^2 & 0 & 0 \\ 0 & \frac{1}{\beta} & 0 \\ 0 & 0 & \frac{1}{\beta} \end{pmatrix}, \tag{18}$$

其中，$\beta \in \mathbb{R}_{>0}$是一个模型参数 [50]。因此，$\mathsf{S}^{\|} = \{\boldsymbol{G}\}$成立。

#### 2.3.1 各向同性模型

此时，解析模型必须明确选择超弹性势的函数关系。虽然某些选择有很强的物理依据 [39], ，但大多数模型都带有启发式性质——而这种人为选择函数关系所导致的灵活性降低，纯粹是因为模型必须具有明确的形式。具体而言，我们考虑各向同性 Neo-Hooke 模型

$$\psi^{\mathrm{nh}}(I_1, I_3) = \frac{1}{2}\left(\mu\,(I_1 - \ln I_3 - 3) + \frac{\lambda}{2}\,(I_3 - \ln I_3 - 1)\right), \quad \mu = \frac{E}{2(1+\nu)}, \quad \lambda = \frac{E\nu}{(1+\nu)(1-2\nu)} \tag{19}$$

其中材料参数（$E$, $\nu$）分别对应杨氏模量和泊松比。由此势函数可推导出第二类皮奥拉-基尔霍夫应力：

$$\boldsymbol{T}^{\mathrm{nh}} = \mu\mathbf{1} + \left(\frac{\lambda}{2} - \frac{2\mu+\lambda}{2I_3}\right)\operatorname{cof}\boldsymbol{C}\,. \tag{20}$$

请注意，例如，对 $I_1$ 的线性依赖是一种相当严格的选择。同样，通过将应力构造为梯度场，满足热力学第二定律；通过使用不变量，满足柯西应力张量的对称性、客观性和材料对称性。随后，通过巧妙地组合材料参数和不变量，满足归一化条件；将$\boldsymbol{C} = \mathbf{1}$, $I_1 = 3$, $I_3 = 1$ 代入式（19）和（20）即可清楚看出这一点。将项 $-\ln I_3$ 与其余不变量结合使用，也满足体积增长条件。最后，所有数学运算都保持了不变量的多凸性。

令人惊讶的是，据作者所知，对于不变量 $I_1$ 和 $I_3$ 的任意物理上可接受的组合，弹性势的非负性，即 $\psi^{\mathrm{nh}}(I_1, I_3) \geq 0$，尚未得到证明。这尤其具有挑战性，因为 $\psi^{\mathrm{nh}}(I_1, I_3)$ 是多凸的，也就是说，它关于 $\boldsymbol{F}$、$\operatorname{cof}\boldsymbol{F}$ 和 $\det\boldsymbol{F}$ 是凸的，但既不单独关于 $\boldsymbol{F}$ 凸，也不关于 $\boldsymbol{C}$ 凸。[5] 具体而言，Neo-Hooke 势关于 $\boldsymbol{F}$ 并非凸函数，因为不变量 $I_3$ 关于 $\boldsymbol{F}$ [52] 并非凸函数。定理 C.5 给出了可压缩Neo-Hooke势非负性的证明。

[5]这与不可压缩情形 [36], 形成对比；在不可压缩情形下，不变量的定义域要受限得多，因为$J \equiv 1$。

### 2.3.2 横观各向同性模型

为了 为描述横观各向同性材料行为，我们考虑解析模式 1

$$\tilde{\psi}^{\mathrm{ti}}(I_1,I_2,I_3,I_4,I_5)=\alpha_1 I_1+\alpha_2 I_2+\delta_1 I_3-\delta_2\ln(\sqrt{I_3})+\eta^*\left(I_4^{\alpha_4}+I_5^{\alpha_4}\right),\quad \eta^*:=\frac{\eta_1}{\alpha_4(\mathrm{tr}\boldsymbol{G})^{\alpha_4}}, \tag{21}$$

正如 Schröder等人 [50] 所提出的。式（21）中给出的这一横向各向同性势函数不满足能量归一化条件，参见式（13）。为纳入这一条件，我们引入稍作调整后的形式

$$\psi^{\mathrm{ti}}(I_1,I_2,I_3,I_4,I_5):=\tilde{\psi}^{\mathrm{ti}}(I_1,I_2,I_3,I_4,I_5)+\tilde{\psi}^{\mathrm{ti,\,energy}} \tag{22}$$

以及归一化项

$$\tilde{\psi}^{\mathrm{ti,\,energy}}:=-\tilde{\psi}^{\mathrm{ti}}(I_1,I_2,I_3,I_4,I_5)\Big|_{\boldsymbol{C}=\boldsymbol{1}}=-\left(3\alpha_1+3\alpha_2+\delta_1+2\frac{\eta_1}{\alpha_4}\right), \tag{23}$$

由此可见，对于任意复应变能函数，满足能量归一化条件都很容易。由于式（23）中的能量归一化项为常数，因此既不会影响原势函数（式（21））的多凸性，也不会影响其对应的第二类皮奥拉-基尔霍夫应力

$$\boldsymbol{T}^{\mathrm{ti}}=2\Big(\alpha_1\boldsymbol{1}+\alpha_2(I_1\boldsymbol{1}-\boldsymbol{C})+\Big(\delta_1 I_3-\frac{\delta_2}{2}\Big)\boldsymbol{C}^{-1}+\alpha_4\eta^* I_4^{\alpha_4-1}\boldsymbol{G}+\alpha_4\eta^* I_5^{\alpha_4-1}\Big(I_5\boldsymbol{C}^{-1}-\mathrm{cof}(\boldsymbol{C})\cdot\boldsymbol{G}\cdot\boldsymbol{C}^{-1}\Big)\Big). \tag{24}$$

因此，该本构模型包含材料参数（$\beta,\alpha_1,\alpha_2,\delta_1,\delta_2,\alpha_4,\eta_1$），其限制条件见 [50]。本文介绍的所有条件均由该解析模型满足。然而，与第2.3.1节不同的是，据作者所知，应变能密度的非负性无法通过解析方法证明，因此只能通过数值方法验证。

最后需要指出，尽管本节讨论的超弹性本构条件具有坚实的物理和数学依据，但大多数解析模型所采用的受限函数关系却并非如此。借助物理增强神经网络，可以绕过这一局限。

# 3 物理增强神经网络本构模型

上一节讨论了超弹性的本构条件。现在，我们不再像第2.3.1节和第2.3.2节那样为势函数选择解析表达式，而是旨在利用前馈神经网络出色的逼近性能 [1, 28]。下文将介绍整体 PANN 模型。除能量非负性外，该模型能够严格满足后续提出的所有条件，见第2.2节。

## 3.1 基本条件

通过将应力构造为由前馈神经网络直接预测的势函数的梯度，模型的**热力学一致性** 得以满足。如果进一步将势函数表示为不变量的函数，则会自动得到**对称柯西应力张量**[6]，并满足**客观性** 和**材料对称性** 条件。此时，神经网络已经满足基本

[6]需要注意的是，直接将 $\boldsymbol{F}$ 映射到 $\boldsymbol{P}$ 的前馈神经网络，并不一定能得到能够描述$\boldsymbol{\sigma}=\boldsymbol{\sigma}^T$的弹性模型。对于将 $\boldsymbol{F}$ 映射到 $\psi$的模型，同样存在这一限制（$\boldsymbol{F}$）。例如，Naumann 和 Ihlemann [42], 研究了一种违反可积性条件的拟弹性模型，因此应力不再是梯度场。

条件，这是通过构造实现的，即满足热力学一致性、客观性、材料对称性以及应力张量的对称性。

我们假设，各向异性超弹性材料行为可以通过一组不可约且相互独立的不变量 $\mathrm{I} := (I_1, \ldots\ldots, I_m)$ 来描述，其中 $I_\beta(\boldsymbol{C}, \mathrm{S}^{\square})$，$\beta \in \mathbb{N}_{\leq m}$。然而，可能有必要将更多不变量 $I^*_\gamma(\mathrm{I})$，$\gamma \in \mathbb{N}_{\leq A}$（其中含有 $A \in \mathbb{N}_{\geq 0}$）纳入预测势能的自变量列表，例如，以满足额外的物理条件或提高预测的近似质量 [21, 26]。将扩展的不变量集合 $\mathrm{I}^* := (I_1, \ldots\ldots, I_m, I^*_1, \ldots\ldots, I^*_A)$ 作为标量值输出前馈神经网络的输入，并将网络输出作为超弹性势 $\psi(\mathrm{I}^*)$，所得模型的灵活性远超解析表达式。此外，在许多实际应用中，将网络架构限制为仅包含一个隐藏层且其中有 $N^{\mathrm{NN}}$ 个神经元就已足够。采用激活函数 $\mathrm{F}: \mathbb{R} \to \mathbb{R}$（该函数需二阶连续可微），仅使用一个隐藏层的基于不变量的势函数最简单的形式为

$$\psi^{\mathrm{NN},\square}(\boldsymbol{\mathcal{I}}^*) := \sum_{\alpha=1}^{N^{\mathrm{NN}}} W_\alpha \mathscr{F}\left(\sum_{\beta=1}^{m} w_{\alpha\beta} I_\beta + \sum_{\gamma=1}^{A} w^*_{\alpha\gamma} I^*_\gamma + b_\alpha\right), \tag{25}$$

其中，$W_\alpha$、$w_{\alpha\beta}$、$w^*_{\alpha\gamma}$和 $b_\alpha$分别表示权重和偏置值，它们共同构成参数集 $\mathrm{P} \in \mathbb{R}^P$，其中 $P \in \mathbb{N}$ 表示参数总数，这些参数将在标定过程中进行优化，以拟合给定的数据集。请注意，本文介绍的方法均不局限于只有一个隐藏层的前馈神经网络，而是可以直接应用于多层网络架构，参见附录 A。不过，为了更好地说明推导过程，此处仅使用一个隐藏层。随后，通过应用

$$\boldsymbol{T}^{\mathrm{NN},\square} = 2\sum_{\alpha=1}^{m} \frac{\partial \psi^{\mathrm{NN},\square}}{\partial I_\alpha} \frac{\partial I_\alpha}{\partial \boldsymbol{C}} + 2\sum_{\gamma=1}^{A}\sum_{\beta=1}^{m} \frac{\partial \psi^{\mathrm{NN},\square}}{\partial I^*_\gamma} \frac{\partial I^*_\gamma}{\partial I_\beta} \frac{\partial I_\beta}{\partial \boldsymbol{C}}\,. \tag{26}$$

这种对输入量的特殊选择——即不变量——是将物理学纳入神经网络的第一种方式。

## 3.2 多凸性

为确保势函数的**多凸性**，必须使用所考虑对称群的多凸不变量。此外，还必须以特定方式调整网络架构。通过采用凸且单调不减的激活函数 F 和非负权重，可以确保整体势函数 $\psi^{\mathrm{NN}}$、$_{\square}(\mathrm{I}^*)$具有多凸性 [26]。此处采用*Softplus*激活函数 $\mathscr{SP}(x) := \log(1+\exp(x)) \in \mathrm{C}^\infty$，该函数是凸且单调不减的，因此可得以下条件

$$W_\alpha, w_{\alpha\beta}, w^*_{\alpha\gamma} \in \mathbb{R}_{\geq 0}, b_\alpha \in \mathbb{R} \quad \forall \alpha \in \mathbb{N}_{\leq N^{\mathrm{NN}}}, \beta \in \mathbb{N}_{\leq m}, \gamma \in \mathbb{N}_{\leq A}\,. \tag{27}$$

**注释 3.1 (输入凸神经网络（ICNN）。**这种特殊类型的前馈神经网络，即输出关于输入参数为凸的前馈神经网络，称为输入凸神经网络 [3]。为了理解这类神经网络为何凸，我们先退一步，考察单变量函数

$$f: \mathbb{R} \to \mathbb{R}, \quad x \mapsto f(x) := (g \circ h)(x)\,, \tag{28}$$

其中 $f$由两个函数 $g$，$h$ 组成：$\mathbb{R} \to \mathbb{R}$。鉴于上述所有函数均为二阶连续可微，$f(x)$ 在 $x$ 中的凸性等价于其二阶导数非负

$$f''(x) = (g'' \circ h)(x)\, h'(x)^2 + (g' \circ h)(x)\, h''(x) \geq 0\,. \tag{29}$$

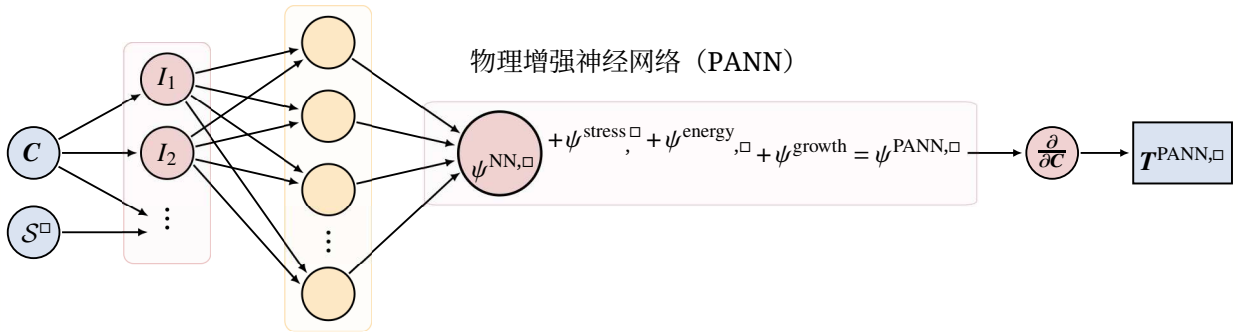


图 2：所考虑材料对称群 □ 对应的 PANN 本构模型示意图。请注意，神经网络的隐藏层（黄色）可以由多层构成。

对此，一个充分但非必要的条件是，最内层函数 $h$ 是凸函数（$h'' \geq 0$），而$g$ 是凸且非递减的（$g'' \geq 0$和 $g' \geq 0$）。这一结论可推广至任意多个函数复合：最内层函数必须是凸函数，而其后的每个函数都必须是凸且非递减的。将其应用于前馈神经网络（FFNN），这类网络可视为多个向量值函数的复合，便可推广得到如下条件：在输入凸神经网络（ICNN）中，第一个隐藏层必须逐节点凸，而后续每一层都必须逐节点凸且非递减，参见 [26, 附录 A]。

应用于超弹性时，输入凸神经网络的架构需要进一步调整：由于不变量是多凸性条件中所定义参数的非线性函数，即（$\boldsymbol{F}$，cof $\boldsymbol{F}$，det $\boldsymbol{F}$），参见式（10），它们是作用于多凸性条件参数上的最内层函数。因此，第一个隐藏层就必须是凸且非递减的。唯一的例外是将 $J = \sqrt{I_3}$用作额外的不变量（而非 $I_3$），因为 $J$ 是多凸性条件的一个参数。因此，作用于 $J$ 的激活函数只需满足凸性，不一定非递减。实际处理中，我们通过在不变量集合中加入额外的不变量 $I_1^* := -2J$ 来考虑这一点；此外，这对于表示负应力也是必不可少的 [26]。请注意，构造多凸神经网络有多种方法，例如 [55] 或 [8]。不过，输入凸神经网络结构简单且灵活性出色，因此是完成这一任务的自然选择。关于多凸神经网络及其明确证明的更全面介绍，请参见 [26]。

### 3.3 生长与归一化条件

最后，还需要在神经网络中加入**生长和归一化条件**，以确保模型具有符合物理规律的应力行为。为此，在式（25）的原始势函数 $\psi^{\mathrm{NN},\square}(\mathrm{I}^*)$ 中加入生长项和归一化项。然后，以 □ 表示所考虑的对称群，整体 PANN 模型可表示为

$$\psi^{\mathrm{PANN},\square}(\boldsymbol{\mathcal{I}}^*) := \psi^{\mathrm{NN},\square}(\boldsymbol{\mathcal{I}}^*) + \psi^{\mathrm{stress},\square}(J, I_4, \ldots, I_m) + \psi^{\mathrm{energy},\square} + \psi^{\mathrm{growth}}(J), \tag{30}$$

以精确的方式满足第2.2节中引入的所有本构条件，但能量非负性除外。即使对于简单的解析模型，例如 Neo-Hooke模型（参见式 (19)），证明能量非负性也并非易事（参见附录 C）。因此，显而易见，要证明相当一般的超弹性势具有非负性是一项艰巨的任务。因此，本文通过数值方法检验PANN模型的非负性（参见第 4 节）；对于各向同性模型，可根据解析考量缩减所需的采样空间（参见定理 C.6）。图 2 以一个隐藏层为例，展示了 PANN模型的整体结构。因此，相应的应力可表示为

如下表达式

$$\boldsymbol{T}^{\mathrm{PANN},\square} = 2\frac{\partial \psi^{\mathrm{PANN},\square}}{\partial \boldsymbol{C}} = \boldsymbol{T}^{\mathrm{NN},\square} + \boldsymbol{T}^{\mathrm{stress},\square} + \boldsymbol{T}^{\mathrm{energy},\square} + \boldsymbol{T}^{\mathrm{growth}} . \tag{31}$$

满足体积**增长条件**的一种方法是使用强制函数。然而，由于输入凸神经网络不一定具有强制性，因此不适合满足该条件，因而需要引入一个解析项

$$\psi^{\mathrm{growth}}(J) := \left(J + \frac{1}{J} - 2\right)^2 \tag{32}$$

该解析项的选取方式应确保不会破坏多凸性。由此得到相应的应力贡献

$$\boldsymbol{T}^{\mathrm{growth}} = 2\left(J + \frac{1}{J} - 2\right)\left(1 - \frac{1}{J^2}\right) J\boldsymbol{C}^{-1} . \tag{33}$$

满足体积增长条件的另一种方法是进一步调整网络架构，参见 [21],不过，使用解析项更为直接。

能量归一化的修正项为

$$\psi^{\mathrm{energy},\square} := -\psi^{\mathrm{NN},\square}(\boldsymbol{\mathcal{I}}^*)\Big|_{\boldsymbol{C}=\boldsymbol{1}} \in \mathbb{R}. \tag{34}$$

由于 $\psi^{\mathrm{energy},\square}$ 是一个常数，因此有 $\boldsymbol{T}^{\mathrm{energy},\square} = \boldsymbol{0}$。再结合记作 $\psi^{\mathrm{stress},\square}$ 的应力归一化项（$J, I_4, \ldots, I_m$），即势函数在未变形状态下的梯度为零，我们确保

$$\psi^{\mathrm{PANN},\square}(\boldsymbol{\mathcal{I}}^*)\Big|_{\boldsymbol{C}=\boldsymbol{1}} = 0 \tag{35}$$

以及

$$\boldsymbol{T}^{\mathrm{PANN},\square}\Big|_{\boldsymbol{C}=\boldsymbol{1}} = 2\left(\sum_{\alpha=1}^{m} \frac{\partial \psi^{\mathrm{PANN},\square}}{\partial I_\alpha}\frac{\partial I_\alpha}{\partial \boldsymbol{C}} + \sum_{\gamma=1}^{A}\sum_{\beta=1}^{m} \frac{\partial \psi^{\mathrm{PANN},\square}}{\partial I^*_\gamma}\frac{\partial I^*_\gamma}{\partial I_\beta}\frac{\partial I_\beta}{\partial \boldsymbol{C}}\right)\Bigg|_{\boldsymbol{C}=\boldsymbol{1}} = \boldsymbol{0} . \tag{36}$$

因此，该势函数仅在未变形状态 $\boldsymbol{C} = \boldsymbol{1}$ 处具有局部极小值；对于此情形，我们将不变量定义为 $I^0_\alpha := I_\alpha(\boldsymbol{C} = \boldsymbol{1})$ 和 $I^{*0}_\beta := I^*_\beta(\boldsymbol{C} = \boldsymbol{1})$。然而，由于 $\psi^{\mathrm{PANN},\square}(\mathrm{I}^*)$ 是多凸的，即关于 $\boldsymbol{F}$、$\mathrm{cof}\,\boldsymbol{F}$ 和 $\det\boldsymbol{F}$ 为凸，但关于 $\boldsymbol{F}$ 或 $\boldsymbol{C}$ 不为凸，因此能量非负性（即 $\psi^{\mathrm{PANN},\square}(\mathrm{I}^*) \geq 0$）并不能自动成立。正如第 2.3.1 节和第 2.3.2 节所述，需要对可容许变形状态进行数值检验，以证明这一条件得到满足，另见第 4 节。

多凸应力修正 $\psi^{\mathrm{stress},\square}(J, I_4, \ldots, I_m)$ 取决于所考虑的对称群，下面将进一步讨论。

### 3.3.1 各向同性归一化项

在各向同性情形下，$\mathrm{I}^* := (I_1, I_2, I_3, I^*_1)$ 的归一化项为

$$\psi^{\mathrm{stress},\circledcirc}(J) := -\mathfrak{n}(J-1) , \tag{37}$$

其中常数

$$\mathfrak{n} := 2\left(\frac{\partial \psi^{\mathrm{NN},\circledcirc}}{\partial I_1} + 2\frac{\partial \psi^{\mathrm{NN},\circledcirc}}{\partial I_2} + \frac{\partial \psi^{\mathrm{NN},\circledcirc}}{\partial I_3} + \frac{\partial \psi^{\mathrm{NN},\circledcirc}}{\partial I^*_1}\frac{\partial I^*_1}{\partial I_3}\right)\Bigg|_{\boldsymbol{C}=\boldsymbol{1}} \in \mathbb{R} \tag{38}$$

是 ICNN 势函数在未变形状态$\boldsymbol{C}=\mathbf{1}$下对不变量的导数的加权和。相应的应力贡献为

$$\boldsymbol{T}^{\text{stress},\circledcirc}=-\mathfrak{n}J\boldsymbol{C}^{-1}\,. \tag{39}$$

并由此可知，未变形状态在构造上是无应力的。

修正项的这一定义包含两层含义：首先，这种方法保留了势函数的多凸性，因为附加项是关于 $J$ 的线性函数，而 $J$ 是多凸性条件参数中包含的一个不变量，参见式（10）。此外，偏导数

$$\left.\frac{\partial I_\alpha}{\partial \boldsymbol{C}}\right|_{\boldsymbol{C}=\mathbf{1}}=\xi\mathbf{1}\quad\text{and}\quad\left.\frac{\partial I_1^*}{\partial \boldsymbol{C}}\right|_{\boldsymbol{C}=\mathbf{1}}=\eta\mathbf{1} \tag{40}$$

各向同性不变量关于 $\boldsymbol{C}$ 的偏导数，在未变形状态下是单位张量 $\mathbf{1}$ 的倍数，对所有 $\alpha\in\mathbb{N}_{\leq 3}$ 均成立，其常数为 $\xi,\ \eta\in\mathbb{R}$。因此，它可以将未变形状态下的应力归一化为零；在式（39）中令 $\boldsymbol{C}=\mathbf{1}$ 即可清楚地看出这一点。更多细节见附录B，其中以解析形式给出了所有必要的张量导数。

### 3.3.2 横观各向同性归一化项

对于满足 $\mathrm{I}^*:=(I_1,\ I_2,\ I_3,\ I_4,\ I_5,\ I_1^*)$ 的横观各向同性情形，附加不变量 $I_4,\ I_5$ 关于 $\boldsymbol{C}$ 的偏导数在未变形状态下不再是单位张量的倍数，而是包含结构张量 $\boldsymbol{G}$，例如：

$$\left.\frac{\partial I_4}{\partial \boldsymbol{C}}\right|_{\boldsymbol{C}=\mathbf{1}}=\boldsymbol{G}\quad\text{and}\quad\left.\frac{\partial I_5}{\partial \boldsymbol{C}}\right|_{\boldsymbol{C}=\mathbf{1}}=I_5\mathbf{1}-\boldsymbol{G}\,. \tag{41}$$

因此，引入修正项

$$\psi^{\text{stress},\parallel}(J,I_4,I_5):=-\mathfrak{o}(J-1)+\mathfrak{p}(I_4-I_4^0)+\mathfrak{q}(I_5-I_5^0) \tag{42}$$

，其中 $I_4^0=I_5^0=\operatorname{tr}\boldsymbol{G}$。同样，常数

$$\mathfrak{o}:=2\left(\frac{\partial\psi^{\text{NN},\parallel}}{\partial I_1}+2\frac{\partial\psi^{\text{NN},\parallel}}{\partial I_2}+\frac{\partial\psi^{\text{NN},\parallel}}{\partial I_3}+\frac{\partial\psi^{\text{NN},\parallel}}{\partial I_1^*}\frac{\partial I_1^*}{\partial I_3}+\frac{\partial\psi^{\text{NN},\parallel}}{\partial I_5}\operatorname{tr}\boldsymbol{G}+\mathfrak{q}\operatorname{tr}\boldsymbol{G}\right)\Bigg|_{\boldsymbol{C}=\mathbf{1}}\in\mathbb{R} \tag{43}$$

是未变形状态 $\boldsymbol{C}=\mathbf{1}$下 ICNN 势函数关于不变量的导数的加权和。此外，记 ReLU函数为 RL，非负常数

$$\mathfrak{p}:=\mathscr{RL}(-x)\in\mathbb{R}_{\geq 0},\quad \mathfrak{q}:=\mathscr{RL}(x)\in\mathbb{R}_{\geq 0} \tag{44}$$

定义为以该参数为自变量

$$x:=\left(\frac{\partial\psi^{\text{NN},\parallel}}{\partial I_4}-\frac{\partial\psi^{\text{NN},\parallel}}{\partial I_5}\right)\Bigg|_{\boldsymbol{C}=\mathbf{1}}. \tag{45}$$

由于 $\mathfrak{p},\mathfrak{q}$ 非负，且 $I_4,I_5$ 具有多凸性，因此修正项同样是多凸的。需要再次强调的是，修正项不对神经网络中隐藏层的数量作任何假设，因此该方法可直接应用于多层网络架构。总的来说，在横观各向同性情形下，归一化项的应力贡献为

$$\boldsymbol{T}^{\text{stress},\parallel}=-\mathfrak{o}J\boldsymbol{C}^{-1}+2\mathfrak{p}\boldsymbol{G}+2\mathfrak{q}\left(I_5\boldsymbol{C}^{-1}-\operatorname{cof}(\boldsymbol{C})\cdot\boldsymbol{G}\cdot\boldsymbol{C}^{-1}\right), \tag{46}$$

其中已纳入附录B中所有必要的张量导数。

请注意，在所有修正项的定义中，尽管 $\psi^{\text{stress},\square}(J, I_4, I_5)$ 和 $\psi^{\text{energy}}$ 取决于在 $\boldsymbol{C} = \mathbf{1}$ 处对 $\psi^{\text{NN},\square}(\mathrm{I}^*)$ 及其偏导数的求值，但并未对 ICNN $\psi^{\text{NN},\square}(\mathrm{I}^*)$ 中隐藏层的数量作出任何假设，并且这些方法可以直接应用于多层网络架构。

**注3.2**。 本节提出的应力归一化项再次体现了同时满足多个力学条件所面临的挑战。在不考虑多凸性的情况下，应力归一化项的构造十分直接。在这种情况下，可以采用诸如

$$\tilde{\psi}^{\text{stress},\square,\text{C}} = -\left.\frac{\partial \psi^{\text{NN},\square}}{\partial \boldsymbol{C}}\right|_{\boldsymbol{C}=\mathbf{1}} : (\boldsymbol{C} - \mathbf{1}) \tag{47}$$

所述的投影方法，如 [10, 式（20）], 另见 [4, 20]和 [57], , 其中采用了式（47）的略作调整版本。遗憾的是，由于 $\boldsymbol{C}$ 的所有分量并非在 $\boldsymbol{F}$ 中都是凸的，因此该方法不是多凸的。此外，由于修正项直接在 $\boldsymbol{C}$ 中构造，因此它不满足材料对称性条件，参见式（ 9 ）。式（ 47 ）的一个多凸对应形式为

$$\tilde{\psi}^{\text{stress},\square,\text{F}} = -\left.\frac{\partial \psi^{\text{NN},\square}}{\partial \boldsymbol{F}}\right|_{\boldsymbol{F}=\mathbf{1}} : (\boldsymbol{F} - \mathbf{1}) \,. \tag{48}$$

然而，由于该方法直接以 $\boldsymbol{F}$ 表述，因此不满足客观性条件、角动量平衡和材料对称性条件，参见第2.2节。类似地，也可以用不变量来表述投影方法。对于本文提出的模型，这对应于

$$\tilde{\psi}^{\text{stress},\square,\text{I}} = -\sum_{\beta=1}^{m} \left.\frac{\partial \psi^{\text{NN},\square}}{\partial I_\beta}\right|_{\boldsymbol{C}=\mathbf{1}} \left(I_\beta - I_\beta^0\right) - \sum_{\gamma=1}^{A} \left.\frac{\partial \psi^{\text{NN},\square}}{\partial I_\gamma^*}\right|_{\boldsymbol{C}=\mathbf{1}} \left(I_\gamma^* - I_\gamma^{*0}\right) \,. \tag{49}$$

该投影方法确保修正势的偏导数

$$\tilde{\psi}^{\text{NN},\square} := \psi^{\text{NN},\square} + \tilde{\psi}^{\text{stress},\square,\text{I}} \tag{50}$$

相对于每个单独不变量，在单位张量 $\boldsymbol{C} = \mathbf{1}$ 处均为零，从而确保参考构形无应力。进而，这同时满足客观性和材料对称性。但同样，基于不变量的投影方法不满足多凸性条件，因为此处所需的减法运算无法保持不变量的多凸性。

与基于不变量的投影方法密切相关的是，可以直接表述超弹性势，使其相对于不变量的梯度在参考构形 [36, 55] 处为零。对于本文提出的模型，这对应于

$$\left.\frac{\partial \psi^{\text{NN},\square}}{\partial I_\alpha}\right|_{\boldsymbol{C}=\mathbf{1}} = 0 \quad \text{and} \quad \left.\frac{\partial \psi^{\text{NN},\square}}{\partial I_\gamma^*}\right|_{\boldsymbol{C}=\mathbf{1}} = 0 \quad \forall \alpha \in \mathbb{N}_{\leq m}, \gamma \in \mathbb{N}_{\leq A} \,. \tag{51}$$

在 [36], 中，这是通过应用诸如

$$f(I_1) = (I_1 - 3)^2 \,, \tag{52}$$

其梯度

$$\frac{\partial f(I_1)}{\partial I_1} = 2(I_1 - 3) \tag{53}$$

在单位张量 $\boldsymbol{C}=\mathbf{1}$，$I_1=3$ 处为零。然而，正如引言中所指出的，[36] 考虑的是不可压缩材料行为，因此可以使用简单的激活函数，例如二次函数，参见式（52）。在可压缩情形下，不变量在单位张量 $\boldsymbol{C}=\mathbf{1}$ 处没有全局最小值，适用的激活函数便会变得相当复杂。构造这类函数尤其具有挑战性，因为除了要在某个特定点处梯度为零之外，函数还应满足其他条件。例如，它应当是凸且单调递增的，至少二阶连续可微，并且斜率不应过大。总的来说，$\alpha\in(0,\ \infty)$ 的一个候选函数是这个相当复杂的函数

$$f(x)=\begin{cases}0 & \text{for } x<0\,,\\ \frac{\alpha}{6}x^3-\frac{1}{12}x^4 & \text{for } 0\le x\le\alpha\,,\\ \frac{\alpha^3}{6}x-\frac{\alpha^4}{12} & \text{for } \alpha<x\,.\end{cases}\tag{54}$$

此外，尽管这种方法适用于构造多凸模型，但它的限制相当严格，因为它甚至无法表示简单的解析模型，例如 Neo-Hooke模型，参见公式 (19) 和参考文献 [55]。在 [36], 中，这个问题通过假设完全不可压缩性，并将静水压力作为额外未知量来规避。然而，对于一般的可压缩情形，这种做法并不可行，这表明此类方法仅适用于不可压缩情形。

横观各向同性的另一种方法是对势函数作如下构造，使得

$$\frac{\partial\psi^{\mathrm{NN},\parallel}}{\partial I_4}\equiv\frac{\partial\psi^{\mathrm{NN},\parallel}}{\partial I_5}\tag{55}$$

对于每个变形状态 $\boldsymbol{C}\in\mathcal{Sym}$，该关系均成立。实际上，公式 (55) 类似于解析本构模型的应力归一化方法，参见 [50, 第 3 节。1], [9, 第 6.4 节。1] 以及公式 (21)。在本文提出的框架中，公式 (55) 将有助于构造应力归一化项，参见公式 (42,45)。这样一来，横观各向同性应力归一化项（参见公式 (42)）将具有一种更简单的形式，类似于各向同性情形（参见公式 (37)）。在神经网络的背景下，可以通过将作用于 $I_4$ 和 $I_5$ 输入的所有权重和激活函数设为相同值来满足公式 (55)。然而，这种做法限制较大，因为它会全局限制 $I_4$ 和 $I_5$中超弹性势的函数依赖。

与刚才讨论的方法相比，本文提出的基于不变量的归一化项仅对本构条件应变能势作出较小、限制性较弱的修改，参见注释3.3，同时通过构造满足所有常见的本构条件。

## 3.4 模型校准

最后，必须针对特定材料的数据对模型进行校准，其中包括结构张量集合，并汇总于$\mathsf{S}^{\square}$中；该集合对应于所考虑的材料对称群 $\square$。在本文中，采用如下形式的数据集

$$\mathcal{D}=\left\{\left({}^1\boldsymbol{C},{}^1\boldsymbol{T}\right),\left({}^2\boldsymbol{C},{}^2\boldsymbol{T}\right),\dots\right\}\tag{56}$$

该数据集由以右柯西-格林变形和第二类皮奥拉-基尔霍夫应力张量表示的应变-应力组构成。为了考察模型的泛化能力，即其对一般载荷工况的预测，必须在标定过程中未见过的数据上对其进行评估。因此，整个数据集D被划分为标定数据集$\mathsf{D}_\mathrm{c}$和测试数据集$\mathsf{D}_\mathrm{t}$，其中$\mathsf{D}_\mathrm{c}\cap\mathsf{D}_\mathrm{t}=\varnothing$。随后，在数据集$\mathsf{D}_\mathrm{c}$上完成模型校准后，可在$\mathsf{D}_\mathrm{t}$上评估其预测结果。只有当模型能够预测测试数据集中包含的载荷工况，并且测试数据集中包含的载荷工况足够一般时，才能认为该模型具有良好的泛化能力，并能预测任意变形下的应力。

模型校准的下一步是选择其超参数，即隐藏层的数量及其包含的节点数量。为此，应选择“足够大”的层数和节点数，使模型能够灵活地表征所考虑的材料行为。

对于一般的神经网络而言，过拟合 [1, 28], 即对标定数据进行过于精确的插值，从而导致对一般数据的预测效果不佳 [1, 28], 是一个问题。然而，正如我们将要展示的，对于本文提出的 PANN，物理学的引入赋予了模型显著的数学结构，使其不易发生过拟合。随后，在模型架构固定的情况下，必须对模型参数 P，即其权重和偏置进行优化。为了在数据集 $\mathrm{D_c}$上校准模型参数 P，将以均方误差定义的损失函数

$$\mathscr{MSE}^{\square}(\boldsymbol{\mathcal{P}}) = \frac{1}{|\mathcal{D}_{\mathrm{c}}|}\sum_{i=1}^{|\mathcal{D}_{\mathrm{c}}|}\left\| {}^{i}\boldsymbol{T} - \boldsymbol{T}^{\mathrm{model},\square}({}^{i}\boldsymbol{C};\boldsymbol{\mathcal{P}})\right\|^2 \tag{57}$$

其中 $|\mathrm{D_c}|$ 表示$\mathrm{D_c}$ 中的元组数量，k·k表示弗罗贝尼乌斯范数。请注意，超弹性势 $\psi^{\mathrm{PANN}}(\mathrm{I}^*)$仅通过其梯度进行标定，即通过相应的应力张量进行标定，这称为Sobolev训练 [60]。在优化过程中，采用*SLSQP*优化器（序列最小二乘规划） [44]。根据所提出的模型，损失函数和优化器不会影响底层物理规律，也可以选择其他方案。所述工作流程使用*Python*、*TensorFlow* 和*SciPy*实现。

**注释3.3**。 在标定过程中，重要的是要预先考虑模型 $\psi^{\mathrm{PANN},\square}(\mathrm{I}^*)$中包含的增长项 $\psi^{\mathrm{growth}}(J)$，参见式（30）；尤其对于包含较大体积压缩的标定数据，该增长项会对模型产生显著影响。另一方面，应力和能量归一化项未必需要纳入标定过程：能量归一化项 $\psi^{\mathrm{energy}}$ 可以在标定后添加到PANN模型中；由于它是常数，因此完全不会影响应力预测。这样一来，即使模型在未包含应力修正项 $\psi^{\mathrm{stress},\square}(J, I_4, \ldots, I_m)$的情况下进行标定，也能够通过合适的数据学习如何近似无应力参考构形 [26]。只要对无应力参考构形的近似足够准确，应力修正所依赖的常数 $\mathfrak{n}, \mathfrak{o}, \mathfrak{p}$ 就会变得非常小，参见式（38、43、44）。此时，修正项 $\psi^{\mathrm{stress},\square}(J, I_4, \ldots, I_m)$及其梯度 $\boldsymbol{T}^{\mathrm{stress},\square}$都近似为零，因此在模型标定后添加应力修正，对整体模型行为的影响微乎其微。虽然通过合适的数据学习，确实可以近似满足归一化条件，但只有将其纳入模型公式（例如通过归一化项），才能使其得到精确满足。

# 4 数值示例

在上一节介绍了 PANN 超弹性本构模型后，本节将通过若干数值示例展示我们所用方法的能力，并将其与物理基础较弱的方法进行比较。我们先从简单应力状态入手，分析插值和外推行为，以及扰动数据下的表现。随后，我们将研究复杂的多轴变形状态。最后，将训练好的 PANN模型应用于有限元模拟中。

## 4.1 简单应力—应变状态

在本小节中，我们比较三种类型的模型在简单应力-应变状态下的插值和外推能力：一种直接将 $\boldsymbol{F}$ 映射到 $\boldsymbol{P}$的网络 $\boldsymbol{P}^{\mathrm{simple}}(\boldsymbol{F})$，一种考虑基本条件的网络 $\psi^{\mathrm{NN},\circledcirc}(\mathrm{I}^*)$，以及 PANN $\psi^{\mathrm{PANN},\circledcirc}(\mathrm{I}^*)$。因此，第一种

模型可表示为

$$P_{kL}^{\text{simple}}(\boldsymbol{F}) := B_{kL} + \sum_{\alpha=1}^{N^{\text{NN}}} W_{\alpha kL}\, \mathcal{SP}\left(\sum_{i=1}^{3}\sum_{J=1}^{3} w_{\alpha iJ} F_{iJ} + b_\alpha\right) \text{ with } B_{kL}, W_{\alpha kL}, w_{\alpha iJ}, b_\alpha \in \mathbb{R}\,, \tag{58}$$

未考虑任何已引入的条件；第二个模型由式（25）定义，满足热力学一致性、 $\boldsymbol{\sigma}$的对称性、客观性和材料对称性；而 PANN 除满足前一个模型的条件外，还纳入了多凸性、增长条件以及能量和应力归一化，见式（30）。

在以下示例中，三个模型的架构均设为一个隐藏层，包含 $N^{\text{NN}} := 4$ 个神经元。此外，所有模型均采用 Softplus 激活函数。考虑基本条件的模型和 PANN 的标定按照第 3.4 节进行。相比之下，$\boldsymbol{F}$-$\boldsymbol{P}$ 模型的训练自然使用了针对 $\boldsymbol{F}$ 和 $\boldsymbol{P}$的数据。此外，该模型的训练采用 Adam 优化器。本研究的训练数据通过各向同性 Neo-Hooke 模型生成，参见式（20）；其中选取常数 $(E,\ \nu) :=$ (1MPa，0。3)。

### 4.1.1 插值行为

为了分析这三个模型的插值行为，我们研究了三组不同的单轴应力状态：理想数据、偏移数据以及含噪数据。

**理想数据** 首先，使用来自沿 $X_1$ 方向进行的单轴拉伸与压缩试验的解析数据训练这三个模型，参见图3(a1)。给定沿 $X_1$ 方向的主伸长 $\lambda_1 \in \mathbb{R}_{>0}$，并计算主伸长 $\lambda_2 \in \mathbb{R}_{>0}$，使应力 $T_{22}^{\circledcirc}, T_{33}^{\circledcirc}$为零，从而使变形张量和应力张量的分量满足

$$(C_{KL})^{\text{uniaxial}} := \begin{pmatrix} \lambda_1^2 & 0 & 0 \\ 0 & \lambda_2^2 & 0 \\ 0 & 0 & \lambda_2^2 \end{pmatrix}, \quad (T_{KL})^{\text{uniaxial},\circledcirc} = \begin{pmatrix} T_{11}^{\circledcirc} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{59}$$

因此，这会形成一种单轴应力状态，这种状态常用于实验研究。模型的训练数据仅包含30个元组（${}^i\boldsymbol{C}, {}^i\boldsymbol{T}^{\circledcirc}$），其中 $0.8 \le \lambda_1 \le 2$。请注意，元组 $\lambda_1 = 1$ 在数据集中出现了两次，因为该点同时归属于压缩范围和拉伸范围。这不会影响 PANN模型 $\psi^{\text{PANN},\circledcirc}(\text{I}^*)$ 的逼近行为，因为该模型在构造时进行了归一化。然而，对于 $\boldsymbol{F}$-$\boldsymbol{P}$模型和基本条件模型 $\psi^{\text{NN},\circledcirc}(\text{I}^*)$，未变形状态在数据集中具有更高的权重。

标定后，三个模型的应力-拉伸曲线 $T_{11}^{\text{simple}}(\lambda_1)$、 $T^{\text{NN}}{}_{,11}{}^{\circledcirc}(\lambda_1)$ 和 $T^{\text{PANN}}{}_{,11}{}^{\circledcirc}(\lambda_1)$ 均可用解析形式表示，并在图3(a1)中与训练数据进行了比较。正如预期的那样，可以看到，所有三个基于神经网络的模型都能完美逼近理想数据集，这一点也从表1给出的均方误差中得到了印证。基本条件模型和 PANN 也能以很高的精度再现能量$\psi^{\circledcirc}(\lambda_1)$，尽管它们并未直接针对能量进行训练，见图3(a2)。因此，融入物理学并不会降低模型预测质量。此外，我们想指出，只有 PANN模型真正满足归一化条件，即 $T^{\text{PANN}}{}_{,11}{}^{\circledcirc}(\lambda_1 = 1) = 0$，且满足得完全精确。

**偏移数据** 下一步是研究模型对非理想数据的灵活性。为此，理想数据集的应力分量按照以下方式进行偏移:

$$(T_{KL})^{\text{offset},\circledcirc} := (T_{KL})^{\text{uniaxial},\circledcirc} + \begin{pmatrix} 100 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{kPa}, \tag{60}$$

因此，标定数据集在未变形状态下不再满足归一化条件，即 $T^{\text{offset}}{}_{,11}{}^{\circledcirc}(\lambda_1 = 1) \neq 0$。由此，PANN模型中所包含的归一化条件如今尤为值得关注。

如图3(b1)所示， ***F*-*P*** 模型 $\boldsymbol{P}^{\text{simple}}(\boldsymbol{F})$再次完美地逼近了训练数据，另见表1中的均方误差。不过，这只有在这个简单模型不了解势函数存在的情况下才有可能，因此它并不是真正的超弹性模型。相比之下，基本条件模型 $\psi^{\text{NN},\circledcirc}(\mathrm{I}^*)$无法在 $\lambda = 1$ 附近很好地再现训练数据，这是因为它不能违反材料对称性，且应力由弹性势的导数决定。然而，对于未变形状态，模型预测出非零应力，这与标准弹性模型的预期相矛盾。

如图3(b1)所示，由于第3节所述的模型方法，只有PANN模型 $\psi^{\text{PANN},\circledcirc}(\mathrm{I}^*)$在模型校准后才严格满足归一化条件，即 $T\text{PANN},_{11}{}^{\circledcirc}(\lambda_1 = 1) = 0$。尽管第3节中的证明已清楚表明这一点，但此处仍展示了这一重要条件是如何通过构造始终得到满足的，即使训练数据存在偏移也是如此。与基本条件模型类似，由于应力归一化和多凸性，PANN模型无法很好地近似 $\lambda_1 = 1$ 附近的数据点——但其近似方式具有物理意义——因此应力-拉伸曲线中出现了一个拐点。在这一区域之外，对数据点的近似

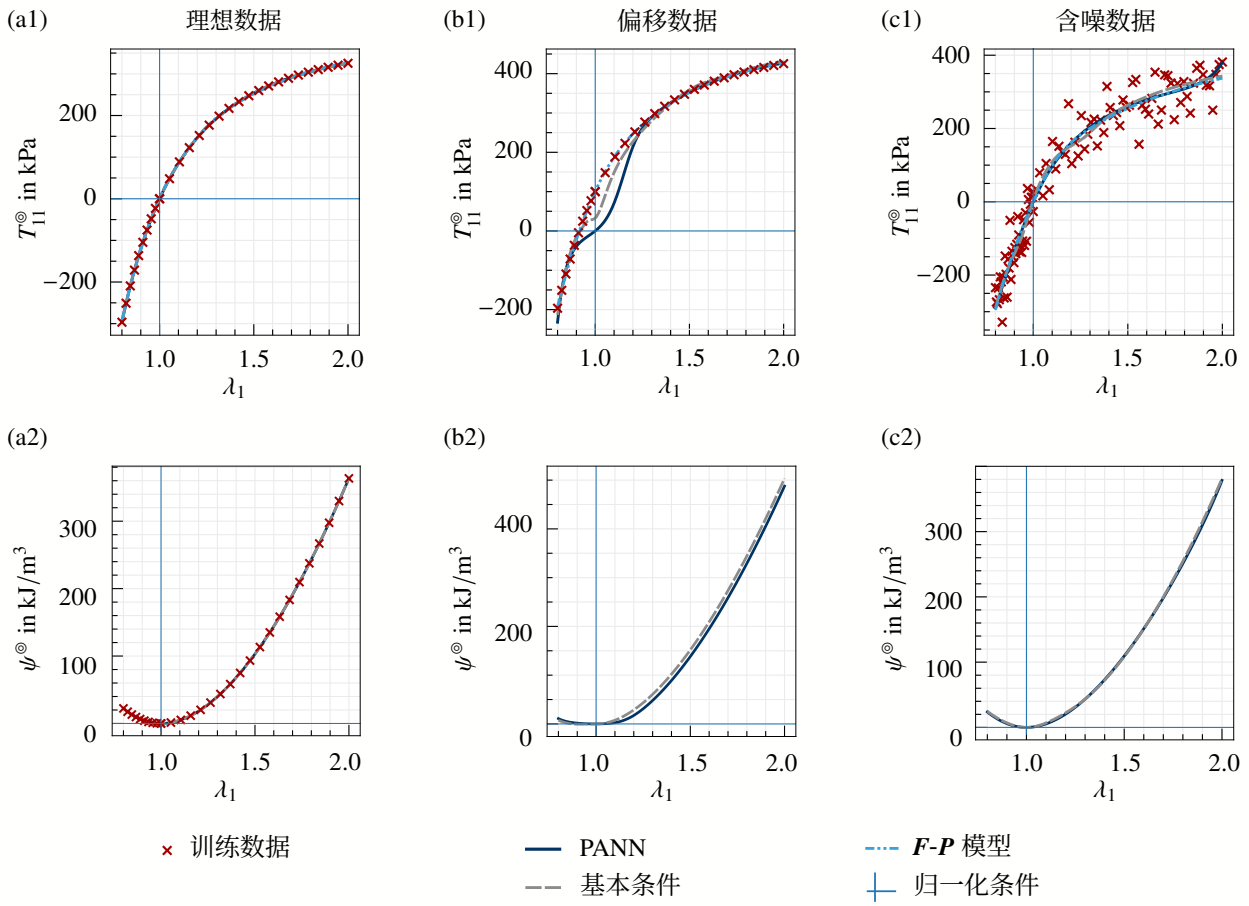


图 3：PANN、满足基本条件的神经网络以及具有一个隐藏层（含有 $N^{\text{NN}} := 4$ 个神经元）的 ***F*-*P***模型所预测的应力-拉伸曲线 $T_{11}^{\circledcirc}(\lambda_1)$。将训练后的模型与 $X_1$ 方向单轴拉伸/压缩试验的数据进行比较，数据分别为：(a1) 理想各向同性 Neo-Hooke 数据，(b1) 带偏移的数据，以及 (c1) 含噪数据。(a2)–(c2) 展示了相应模型的能量。为便于比较，对 $\psi^{\text{NN},\circledcirc}(\mathrm{I}^*)$ 进行了能量归一化，但这不会影响应力预测。

则近乎完美。因此，不应将 $T_{11}$ 在 $\lambda_1$ 中的非凸性误认为违反了多凸性条件。所提出的神经网络模型通过构造具有多凸性，参见注释3.1；此外，多凸性不会对 $\boldsymbol{F}$中 $\boldsymbol{T}$ 的依赖关系施加限制。

由于上述物理限制，基本条件模型和PANN的均方误差相比偏移数据的 ***F-P*** 模型大几个数量级，参见表1。 $\psi^{\mathrm{NN},\circledcirc}(\mathrm{I}^*)$和 $\psi^{\mathrm{PANN},\circledcirc}(\mathrm{I}^*)$给出的弹性能如图3(b2)所示。因此，两个模型之间没有出现显著差异。请注意，训练数据未包含在该图中，因为对于所研究的偏移数据集，没有可用的能量。

**含噪数据** 最后一步研究了三个模型对于含噪数据的逼近行为。为此，将包含100个元组（$^i\boldsymbol{C}$，$^i\boldsymbol{T}^{\circledcirc}$）的更大理想数据集中的应力分量按照下式进行偏移:

$$(T_{KL})^{\mathrm{noisy},\circledcirc} := (T_{KL})^{\mathrm{uniaxial},\circledcirc} + \begin{pmatrix} \xi & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{with} \quad \xi \sim \mathcal{N}(\mu, \sigma^2)\,, \tag{61}$$

其中 $\xi$ 表示　　均值为 $\mu = 0$ 千帕、标准差为 $\sigma = 50$千帕的正态分布噪声　　.

尽管训练数据叠加了很强的高斯噪声，如图3(c1)所示，三个校准模型都给出了看起来物理上合理的结果。更准确地说，预测得到的应力-拉伸曲线单调递增，并且图3(c2)所示的势能 $\psi^{\mathrm{PANN},\circledcirc}(\lambda_1)$和 $\psi^{\mathrm{NN},\circledcirc}(\lambda_1)$关于 $\lambda_1$是凸的，这对于单轴应力而言是预期结果；此外，它们均较为平滑，没有表现出明显的振荡。可以看出，在这种情况下，未变形状态下的应力预测约为零，见图3(c1)。然而，只有PANN模型以严格方式满足所有相关条件，这一点对于训练数据中未包含的应力-应变状态同样适用。该性质的优势将在下一项研究中展现。 ***F-P*** 模型甚至违反了角动量平衡，因为 $\boldsymbol{T}^{\mathrm{simple}}$ 和 $\boldsymbol{\sigma}^{\mathrm{simple}}$ 并不对称。

三个模型的均方误差再次列于表1。由于数据并不完美，与理想情况相比，数值增大了几个数量级。

表1：使用不同单轴应力数据进行训练时， ***F-P*** 模型、满足基本条件的神经网络以及 PANN 根据公式（57）得到的均方误差。相应的应力预测如图3所示。

| 模型 | 理想数据 | 带偏移的数据 | 含噪数据 |
|---|---|---|---|
| ***F-P*** 模型 | 3.14 千帕$^2$ | 2.87 千帕$^2$ | $2.02 \cdot 10^3$ 千帕$^2$ |
| 基本条件神经网络 | 1。$74 \cdot 10^{-4}$ 千帕$^2$ | $1.22 \cdot 10^3$ 千帕$^2$ | 2。$03 \cdot 10^3$ 千帕$^2$ |
| PANN模型 | $5.92 \cdot 10^{-5}$ 千帕$^2$ | $2.77 \cdot 10^3$ 千帕$^2$ | $2.02 \cdot 10^3$ 千帕$^2$ |

### 4.1.2 外推行为

现在，通过考虑三种载荷工况：单轴拉伸/压缩、双轴拉伸/压缩以及简单剪切，来分析所比较模型的外推行为。对于 $\boldsymbol{P}^{\mathrm{simple}}(\boldsymbol{F})$、$\psi^{\mathrm{NN},\circledcirc}(\mathrm{I}^*)$ 和 $\psi^{\mathrm{PANN},\circledcirc}(\mathrm{I}^*)$ 的训练，这里仅使用了单轴应力状态及其在狭窄范围 $0.8 \le \lambda_1 \le 1.1$ 内对应的变形，参见式 (59)；这些数据存储在数据集 $\mathrm{D}^{\mathrm{uniaxial},\circledcirc}$ with $|\mathrm{D}^{\mathrm{uniaxial},\circledcirc}| = 15$ 中。同样需要注意的是，数据集中 $\lambda_1 = 1$ 对应的元组出现了两次，因为该点同时归属于压缩范围和拉伸范围。鉴于 PANN模型 $\psi^{\mathrm{PANN},\circledcirc}(\mathrm{I}^*)$ 在构造上已进行了归一化，这不会影响其逼近行为。相比之下，对于 ***F-P*** 模型和基本条件模型 $\psi^{\mathrm{NN},\circledcirc}(\mathrm{I}^*)$，数据集对未变形状态赋予了更高的权重。

**单轴应力状态** 如图 4(a1) 所示，所有模型再次完美地逼近了训练数据，这一点也可以从表 2 给出的均方误差中看出。然而， ***F-P*** 模型一旦需要进行外推，便立即失效。相比之下，模型 $\psi^{\mathrm{NN},\circledcirc}(\mathrm{I}^{*})$ 在拉伸达到 $\lambda_1 = 4$ 之前的外推效果非常好；在作者看来，这着实令人惊叹。同样，图 4(a2) 中所示的相关能量也得到了很好的再现。因此，基于不变量的方法 [13, 21, 26, 27, 33, 35, 36, 55]（这类方法逼近的是能量，而非直接逼近应力）的优势在此尤为明显。PANN模型的结果更加令人印象深刻：尽管需要外推，它仍能几乎完美地再现数据。外推数据的均方误差也印证了这一点：PANN模型的均方误差比基本条件模型低两个数量级，参见表 2。观察到的改进源于将额外的物理原理，即多凸性、增长条件以及能量和应力归一化，引入了 PANN模型。

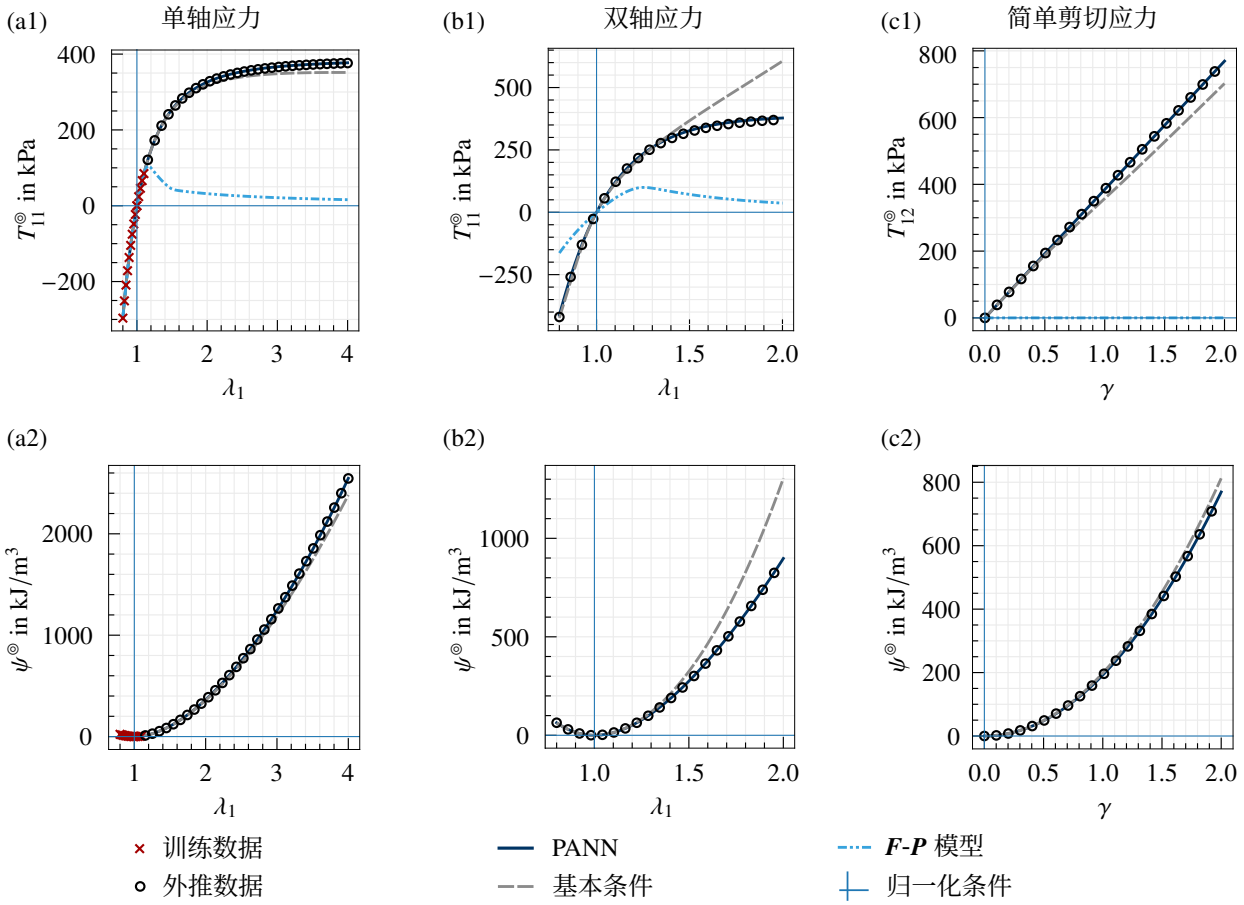


图 4：PANN、满足基本条件的神经网络，以及具有一个隐藏层且包含 $N^{\mathrm{NN}} := 4$ 个神经元的 ***F-P*** 模型所预测的应力-拉伸曲线 $T_{KL}^{\circledcirc}(\lambda_1)$ 和 $T_{KL}^{\circledcirc}(\gamma)$。图中展示了使用存储在 $\mathrm{D}^{\mathrm{uniaxial},\circledcirc}$ 中的单轴应力数据训练的模型，针对以下情况进行的外推结果： (a1) 单轴拉伸/压缩试验，(b1) 双轴拉伸/压缩试验，以及 (c1) 简单剪切试验。对应的能量见 (a2)–(c2)。

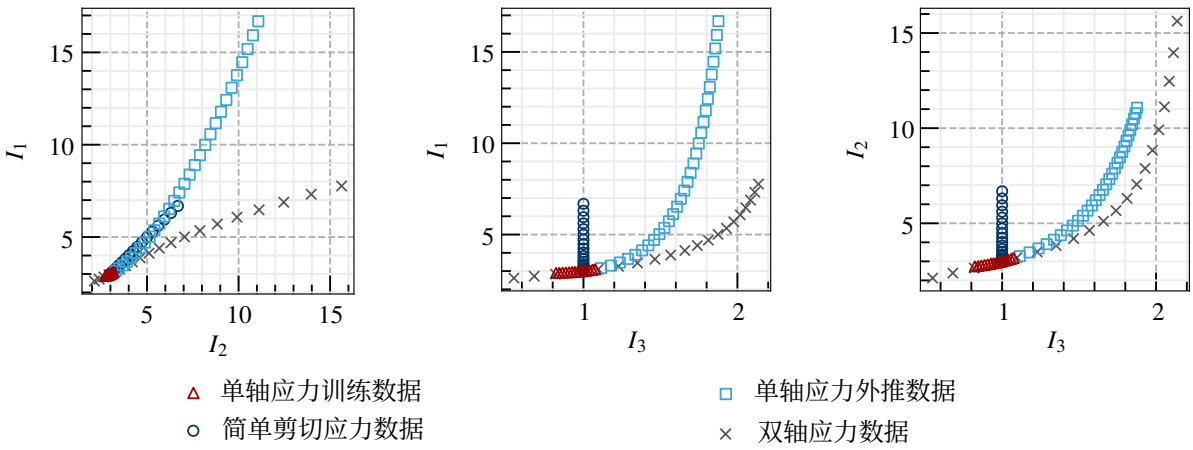


图5：单轴拉伸/压缩试验、双轴拉伸/压缩试验和简单剪切试验中变形状态在各向同性不变量空间中的位置。图中显示了截面 $I_1$-$I_2$、 $I_1$-$I_3$ 和 $I_2$-$I_3$。

**双轴应力状态和简单剪切**接下来，我们要评估模型在必须对完全未知的载荷工况进行外推时的表现，即双轴应力状态

$$(C_{KL})^{\text{biaxial}} := \begin{pmatrix} \lambda_1^2 & 0 & 0 \\ 0 & \lambda_1^2 & 0 \\ 0 & 0 & \lambda_2^2 \end{pmatrix}, \quad (T_{KL})^{\text{biaxial},\circledcirc} = \begin{pmatrix} T_{11}^{\circledcirc} & 0 & 0 \\ 0 & T_{11}^{\circledcirc} & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{62}$$

和简单剪切

$$(C_{KL})^{\text{shear}} := \begin{pmatrix} 1 & \gamma & 0 \\ \gamma & \gamma^2+1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{63}$$

其中 $0 \le \gamma \le 2$ 表示剪切。应力预测结果见图4(b1)和(c1)。同样，简单方法 $\boldsymbol{P}^{\text{simple}}(\boldsymbol{F})$ 对未见数据完全失效。在简单剪切加载情况下，模型甚至预测出完全不合理的剪切应力 $T\text{simple}_{12} \approx 0$，因为模型在训练过程中学到应力张量的对角元素为零。相比之下，基本条件模型 $\psi^{\text{NN},\circledcirc}(\text{I}^*)$ 仍能进行相当好的物理上合理的外推，但在双轴应力加载中，当 $\lambda_1 > 1.4$ 时出现明显偏差。简单剪切也有类似结果，从 $\gamma \approx 0.8$ 开始出现明显偏差。能量近似也呈现相同的情况，结果见图4(b2)和(c2)。令人惊讶的是，尽管这里需要进行完全外推，PANN模型仍能针对双轴和简单剪切几乎完美地预测数据。表 2 中给出的均方误差也清楚地表明，与 $\boldsymbol{P}^{\text{simple}}(\boldsymbol{F})$ 和 $\psi^{\text{NN},\circledcirc}(\text{I}^*)$ 相比，PANN模型 $\psi^{\text{PANN},\circledcirc}(\text{I}^*)$ 的外推能力显著提升。为说明针对所考察的测试案例，各模型必须在哪些范围内进行外推，图5在不变量空间中展示了所考察的变形状态。

因此，综上所述，本研究的发现表明，对基于神经网络的模型进行调整，使其满足任意载荷下的物理条件，并不一定能改善对训练数据的近似效果，至少对于所考察的简单载荷工况而言是如此；但这确实能够显著提升其外推能力。

## 4.2 复杂多轴应力-应变状态

现在，我们已经展示了基于不变量能量的神经网络方法的优势，因此接下来只分析这一类模型。由此，我们关注的是，严格满足物理原理

表2：根据公式（57），使用 ***F*-*P*** 模型、满足基本条件的神经网络以及PANN模型，对单轴拉伸/压缩（$0.8 \leq \lambda_1 \leq 1.1$）训练以及外推至未知状态所取得的均方误差。

| 模型 | 训练单轴应力 | 外推单轴应力 | 双轴应力 | 简单剪切应力 |
|---|---|---|---|---|
| ***F*-*P*** 模型 | $1.95 \cdot 10^{-1}$ 千帕$^2$ | $9.43 \cdot 10^{4}$ 千帕$^2$ | $1.35 \cdot 10^{5}$ 千帕$^2$ | $8.06 \cdot 10^{5}$ 千帕$^2$ |
| 基本条件神经网络 | $1。30 \cdot 10^{-4}$ 千帕$^2$ | $1.26 \cdot 10^{5}$ 千帕$^2$ | $6.24 \cdot 10^{5}$ 千帕$^2$ | $5.65 \cdot 10^{4}$ 千帕$^2$ |
| PANN模型 | $3.91 \cdot 10^{-5}$ 千帕$^2$ | $6.21 \cdot 10^{2}$ 千帕$^2$ | $4.11 \cdot 10^{3}$ 千帕$^2$ | $1.58 \cdot 10^{-5}$ 千帕$^2$ |

第2.2节中引入的物理原理是否可能过于严格，以致模型不够灵活，无法充分近似非线性材料行为所产生的复杂多轴应力-应变状态。由此，满足条件的四种不同架构 $\psi^{(\mathrm{i})}{}_{,\square}(\mathrm{I}^*)$, $\psi^{(\mathrm{ii})}{}_{,\square}(\mathrm{I}^*)$, $\psi^{(\mathrm{iii})}{}_{,\square}(\mathrm{I}^*)$ 以及 $\psi^{(\mathrm{iv})}{}_{,\square}(\mathrm{I}^*)$

(i) 基本条件，即热力学一致性、对称的 $\boldsymbol{\sigma}$、客观性和材料对称性，

(ii) (i) 中的条件 +多凸性，

(iii) (ii) 中的条件 +增长条件，以及

（iv）将（iii）中的条件 +能量和应力归一化

分别针对各向同性行为和横观各向同性行为进行比较。请注意，第一个和第四个模型分别等同于 $\psi^{\mathrm{NN}}{}_{,\square}(\mathrm{I}^*)$和 $\psi^{\mathrm{PANN}}{}_{,\square}(\mathrm{I}^*)$。

### 4.2.1 训练数据的生成

首先，需要获取用于训练神经网络的数据基。在缺乏真实实验的情况下，此处通过数值方法生成。为此，在有限元仿真中对一个虚拟试样施加单轴拉伸试验，见图6(a)和(b)。通过对该结构施加单轴拉伸试验，结构中的孔洞将导致有限元网格中每个积分点处出现相当不均匀的变形梯度。该仿真的目标是生成这类一般性的变形梯度，随后利用第2.3节介绍的解析模型生成数据。此处，试样的几何尺寸设定为 $L_{x_1} \times L_{x_2} \times L_{x_3} = (100 \times 100 \times 5)$ mm。在单轴拉伸试验中，施加最大值为$\hat{u} \leq 40$ mm的位移边界条件，即在平面 $X_1 = L_{x_1}$内满足 $u_1 = \hat{u}$，并在有限元仿真的每个增量中线性增加。相反，在平面 $X_1 = 0$内满足 $u_1 = 0$。在点$(0,\ L_{x_2},\ 0)$处，还规定了位移 $u_2 = 0$；由于点$(0,\ L_{x_2},\ -L_{x_3})$处的固定支承，位移受到约束，即满足 $\boldsymbol{u} = \boldsymbol{0}$，以防止刚体运动。几何生成和网格划分采用工具*Gmsh* [15]实现。

在这一虚拟实验中，在有限元的积分点处经过 30 个增量，收集所需的数据集 $\mathrm{D}^{\circledcirc}$，由元组 $\mathrm{D}_i^{\circledcirc} := ({}^i\boldsymbol{C},\ {}^i\boldsymbol{T}^{\circledcirc})$ 组成。为展示所提出的基于神经网络的方法的能力，选用非线性应力-应变关系 (20) 来描述样本材料的各向同性本构行为。此处，解析模型的材料参数根据表 3 选取。

遵循 Kalina 等人 [22] 的工作，并指定相对容差 $\eta := 1\%$，随后根据对应变形状态 ${}^i\boldsymbol{C}$的不变量（${}^iI_1,\ {}^iI_2,\ {}^iI_3$） $\in \mathbb{R}^3$对数据集进行筛选，得到缩减数据集 $\mathrm{D}^{\mathrm{red},\circledcirc}$，其中包含 $|\mathrm{D}^{\mathrm{red},\circledcirc}| = 963$，用于 PANN 的标定过程。该流程利用了这样一个事实：所引入的模型位于不变量空间，而非变形空间。

最后，可以获得不同的变形状态；对于各向同性情形，这些状态连同 Neo-Hooke模型的相应应力存储在数据集 $\mathrm{D}^{\mathrm{red},\circledcirc}$ 中。对于优选方向平行于 $X_1$ 方向的横观各向同性情形，选取相同的变形状态作为数据基，并

将其连同数据集 $\mathrm{D}^{\mathrm{red},\,\|}$ 中 Schröder模型（24）的相应应力存储，其中 $|\mathrm{D}^{\mathrm{red},\,\|}| = 963$ 亦然。解析横观各向同性模型所选取的参数也列于表 3 中。

表 3：Schröder等人提出的各向同性Neo-Hooke模型和横观各向同性解析模型的材料参数 [50]。这两个模型分别见式 (20) 和式 (24)。

| $E$ | $\nu$ | $\beta$ | $\alpha_1$ | $\alpha_2$ | $\delta_1$ | $\delta_2$ | $\alpha_4$ | $\eta_1$ |
|---|---|---|---|---|---|---|---|---|
| $10^3$ 千帕 | 0.3 | 2 | 8 千帕 | 0 千帕 | 10 千帕 | 56 千帕 | 2 | 10 千帕 |

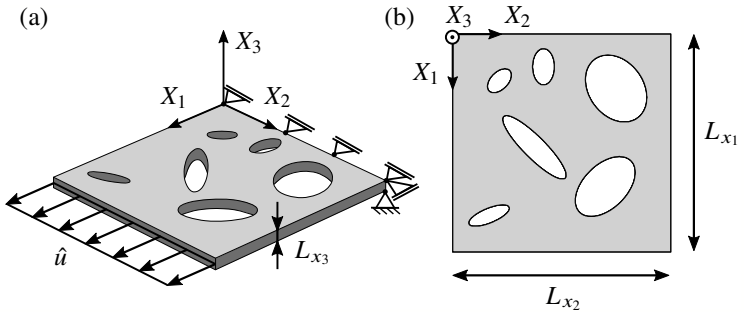


图6：用于数据生成的单轴拉伸试验：(a) 三维非均质试样及施加的边界条件，其中规定的位移为 $\hat{u}\,\boldsymbol{e}_1$；(b) 试样几何形状的俯视图，尺寸为 $L_{x_1} \times L_{x_2} \times L_{x_3} = (100 \times 100 \times 5)$ 毫米。

### 4.2.2 总体预测质量

首先，对于缩减数据集 $\mathrm{D}^{\mathrm{red},\square}$，我们将比较满足不同物理约束的训练神经网络 (i)–(iv) 的总体预测质量。这些物理约束已在第2.2节中介绍，比较对象包括各向同性情形和横观各向同性情形。对于横观各向同性情形，结构张量 $\boldsymbol{G}$ 中的 $\beta = 2$ 取自式 (18)。在以下示例中，网络架构设为包含一个隐藏层，其中有 $N^{\mathrm{NN}} := 8$ 个神经元。基于神经网络的模型针对数据集 $\mathrm{D}^{\mathrm{red},\square}$进行训练，并将数据随机划分为标定数据（70%）和测试数据（30%），且仅进行一次划分。在训练过程中，根据第3.4节确定权重和偏置值 $W_\alpha, b_\alpha, w_{\alpha\beta}$以及 $w^*_{\alpha\gamma}$。在一次训练轮次中，相应的神经网络训练30次，并在结束时保存所达到的均方误差 (57) 最低的最佳训练状态的参数，参见第3.4节 [22]。[7] 为了评估训练神经网络的逼近行为，我们计算相对误差度量

$$\varepsilon^{\square} := \frac{\max\limits_{i \in \mathbb{N}_{\leq |\mathcal{D}^{\square}|}} \left\| {}^{i}\boldsymbol{T}^{\square} - {}^{i}\boldsymbol{T}^{\mathrm{model},\square} \right\|}{\max\limits_{j \in \mathbb{N}_{\leq |\mathcal{D}^{\square}|}} \| {}^{j}\boldsymbol{T}^{\square} \|} \tag{64}$$

对于第二类皮奥拉-基尔霍夫应力的弗罗贝尼乌斯范数 $\mathrm{k}\boldsymbol{T}\mathrm{k}$。为排除随机效应，我们进行了统计研究，即针对每个模型共进行了 300 轮训练。由于结果并未显示误差度量 $\varepsilon^{\square}$ 服从均匀分布，因此此处不假设任何潜在分布，详见附录 D 中给出的直方图。因此，我们使用中位数 $\varepsilon^{\mathrm{med}}$、$\square$ 和分位数来比较各模型的插值质量。图 7 展示了上述研究的结果。

[7] 由于损失函数中存在局部极小值，此处采用的优化过程取决于权重和偏置的初始值。因此，为了克服这一问题，网络会经过多次训练，参见 Kalina 等人 [21, 22]。

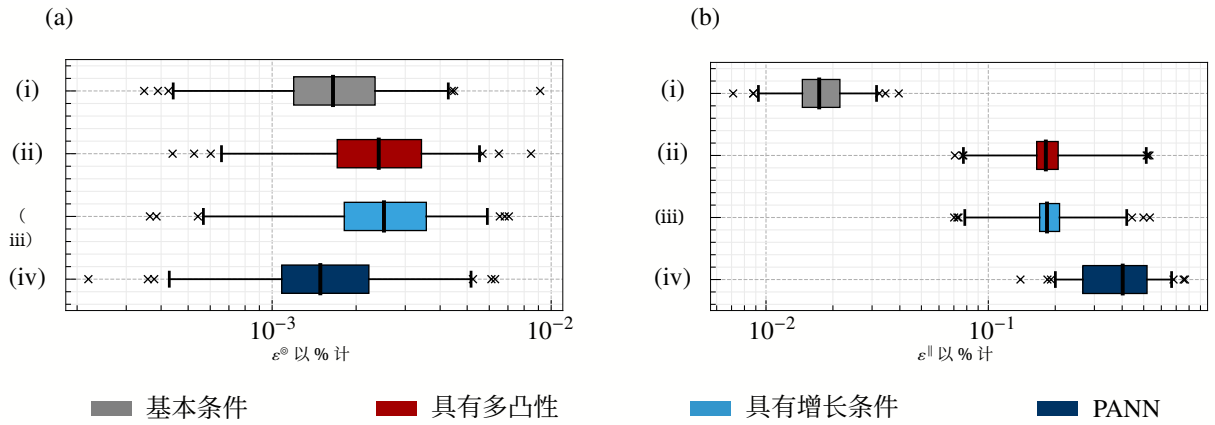


图 7：缩减数据集 $\mathrm{D}^{\mathrm{red},\square}$ 的相对误差度量$\varepsilon^{\square}$（根据式 (64)）箱线图，显示中位数 $\varepsilon^{\mathrm{med},\square}$、第 25 和第 75 百分位数，以及第 1 和第 99 百分位数；对象为 (a) 各向同性神经网络或 (b) 横观各向同性神经网络，分别满足 (i) 基本条件、(ii) 多凸性、(iii) 增长条件和多凸性，以及 (iv) 满足包括归一化在内所有条件的 PANN。

**各向同性模型**关于神经网络在各向同性情形下的应力近似，所有模型都取得了极佳的预测质量，误差中位数为 $\varepsilon^{\mathrm{med},\circledcirc} < 0.003\ \%$，参见图 7(a)。可以看出，与仅满足基本条件的架构 $\psi^{(\mathrm{i}),\circledcirc}(\mathrm{I}^*)$ 相比，考虑多凸性的 $\psi^{(\mathrm{ii}),\circledcirc}(\mathrm{I}^*)$ 的近似质量有所下降。因此，限制权重为正会降低神经网络在统计意义上的近似质量。不过，如前所述，误差仍然极低。若进一步加入增长条件，则该神经网络（记为 $\psi^{(\mathrm{iii}),\circledcirc}(\mathrm{I}^*)$，满足基本条件 + 多凸性 + 增长条件）的误差仍处于相近范围。令人惊讶的是，最终加入归一化条件后，$\psi^{(\mathrm{iv}),\circledcirc}(\mathrm{I}^*)$ 得到的分布与第一个模型，即基本条件模型，非常相似。总而言之，这项研究表明，将超弹性的所有常见物理原理加入基于神经网络的模型，并不会使其在各向同性情形下的近似效果变差。尽管如此，为考虑多凸性而施加正权重等条件，确实会降低神经网络的变异性。

最后，针对弹性能的非负条件，我们进行了数值检验。为仅扫描物理上可接受的变形状态，我们改变主伸长 $\lambda \in \mathbb{R}_{>0}$，并将$\boldsymbol{C}^{\mathrm{diag}}(\lambda) := \lambda^2\mathbf{1}$ 计算为球形张量；根据附录 C 的定理 C.6，这对于各向同性已足够。随后计算不变量并确定相应的能量。在 $1/10 \le \lambda \le 10$的范围内，数值检验仅检测到正能量；例如，对于相对误差为 $\varepsilon^{\circledcirc}$、最接近中位数 $\varepsilon^{\mathrm{med},\circledcirc}$的训练后的各向同性 PANN，情况亦是如此。

**横向各向同性模型** 关于图 7(b) 所示横观各向同性情形下神经网络的应力近似，与各向同性情形相比，即使是最佳模型 $\psi^{(\mathrm{i}),\parallel}(\mathrm{I}^*)$，误差也高了一个数量级。这可以用参考模型 (21) 的复杂度增加来解释。不过，误差的中位数为 $\varepsilon^{\mathrm{med},\parallel} < 0.4\ \%$，所有模型仍然达到了较高的预测质量。现在对四种模型 (i)–(iv) 进行比较。采用考虑多凸性的模型 $\psi^{(\mathrm{ii}),\parallel}(\mathrm{I}^*)$ 后，与满足基本条件的架构 $\psi^{(\mathrm{i}),\parallel}(\mathrm{I}^*)$ 相比，误差增加了一个数量级。因此，对正权重的限制显著降低了神经网络的近似质量。进一步加入增长条件后，该神经网络的误差（记为 $\psi^{(\mathrm{iii}),\parallel}(\mathrm{I}^*)$）仍处于相近范围。最后，加入归一化条件后，模型 $\psi^{(\mathrm{iv}),\parallel}(\mathrm{I}^*)$ 得到的误差分布再次向略大的数值偏移。

我们想强调的是，这些模型仅基于一个包含 $N^{\mathrm{NN}} = 8$ 个神经元的隐藏层。

通过调整网络架构（例如使用多个隐藏层或增加神经元数量），可以轻松提高基于神经网络的模型的近似质量，从而充分表征复杂的材料行为。总而言之，对于横观各向同性情形，将超弹性的所有常见物理原理纳入基于神经网络的模型，会使预测质量下降约一个数量级。尽管如此，误差仍然非常小，因此无论如何都应选择 PANN 模型，尤其是考虑到其出色的外推能力。

关于弹性能的非负条件，也针对横观各向同性进行了数值检验。同样，为了仅扫描物理上容许的变形状态，需要改变主伸长$\lambda_1$、$\lambda_2$、$\lambda_3 \in \mathbb{R}_{>0}$，并计算对角张量 $\boldsymbol{C}^{\mathrm{diag}}(\lambda_1、\lambda_2、\lambda_3) = \lambda_1^2 \boldsymbol{e}_1 \otimes \boldsymbol{e}_1 + \lambda_2^2 \boldsymbol{e}_2 \otimes \boldsymbol{e}_2 + \lambda_3^2 \boldsymbol{e}_3 \otimes \boldsymbol{e}_3$。此外，还必须考虑这些状态绕优选方向 $X_1$ 的旋转。因此，我们最终得到一个用于采样变形的五参数空间：

$$\boldsymbol{C}^{\parallel}(\lambda_1,\lambda_2,\lambda_3,\varphi_2,\varphi_3) = \boldsymbol{R}(\varphi_2,\varphi_3) \cdot \boldsymbol{C}^{\mathrm{diag}}(\lambda_1,\lambda_2,\lambda_3) \cdot \boldsymbol{R}^{T}(\varphi_2,\varphi_3)\ . \tag{65}$$

在上式中，$\boldsymbol{R}(\varphi_2、\varphi_3) = \boldsymbol{R}_{x_2}(\varphi_2) \cdot \boldsymbol{R}_{x_3}(\varphi_3) \in \mathcal{SO}(3)$ 是一个旋转张量，其中 $\boldsymbol{R}_{x_2}(\varphi_2)$ 和 $\boldsymbol{R}_{x_2}(\varphi_3)$ 分别表示绕 $X_2$轴和 $X_3$轴的旋转。结合式 (65) 和式 (18) 中给出的结构张量，可计算不变量并确定相应的能量。在 $1/10 \le \lambda_\alpha \le 10$ 至 $0 \le \varphi_\beta \le \pi/2$ 的范围内，对于 $\alpha \in \mathbb{N}_{\le 3}$、$\beta \in \{2、3\}$，数值检验仅检测到模型 (21) 和训练得到的横观各向同性 PANN 具有正能量；其相对误差为 $\varepsilon^{\parallel}$，该误差最接近中位数$_{\varepsilon^{\mathrm{med},\parallel}}$，这里仅作示例说明。

### 4.2.3 经标定的 PANN 在有限元仿真中的应用

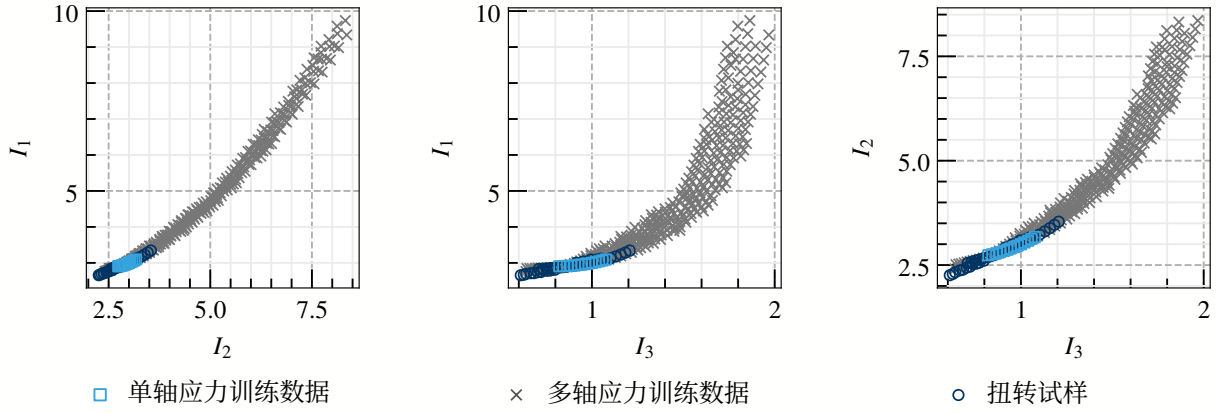


图8：单轴应力-应变数据 $\mathrm{D}^{\mathrm{uniaxial},\circledcirc}$、多轴应力-应变数据 $\mathrm{D}^{\mathrm{red},\circledcirc}$ 以及扭转试样的变形状态在不变量空间中的位置。其中，扭转试样通过指定 $\hat{}\ \phi = 45°$ 的畸变进行加载，参见图 9(a)。图中显示了截面 $I_1$-$I_2$、$I_1$-$I_3$ 和 $I_2$-$I_3$。

为了证明基于不变量的神经网络适用于复杂形状试样和组件的数值仿真，下面将其与采用各向同性 Neo-Hooke 模型 (20) 生成的参考结果进行比较，该模型已用于标定。具体而言，满足基本条件的模型和 PANN 分别在两种不同场景下进行分析：一次使用多轴应力-应变状态 $\mathrm{D}^{\mathrm{red},\circledcirc}$ 进行训练，另一次仅使用单轴应力-应变状态 $\mathrm{D}^{\mathrm{uniaxial},\circledcirc}$进行训练，参见第 4.1.2 节和第 4.2.2 节。在第一种情况下，从 300 个训练模型中选取与中位数 $\varepsilon^{\mathrm{med},\circledcirc}$ 最接近的一个模型。

作为一个测试示例，考虑一个三维问题：棱柱形试样的扭转，参见图 9(a) 中的示意图。因此，鉴于参考本构定律 $\psi^{\mathrm{nh}}(I_1, I_3)$ 的非线性特征，该示例是验证和探索这两种方法潜在局限性的良好机会。在验证

设置中，通过在平面 $\max_{\boldsymbol{x}\in\mathcal{B}_0} X_1$ 内指定 $\hat{\phi} = 45°$ 为扭转角，对带有一个圆孔的三维扭转载荷试样施加载荷。此外，在该平面内施加位移 $u_1 = 0$，以防止翘曲。为避免刚体运动，在平面 $\min_{\boldsymbol{x}\in\mathcal{B}_0} X_1$ 内将位移完全固定，即满足 $\boldsymbol{u} = \boldsymbol{0}$。基本条件模型 (25) 和 PANN (30) 均已在有限元工具箱 *FEniCS* [2, 38] 中实现。其中，方程 (26) 和 (31) 给出的应力关系以及材料

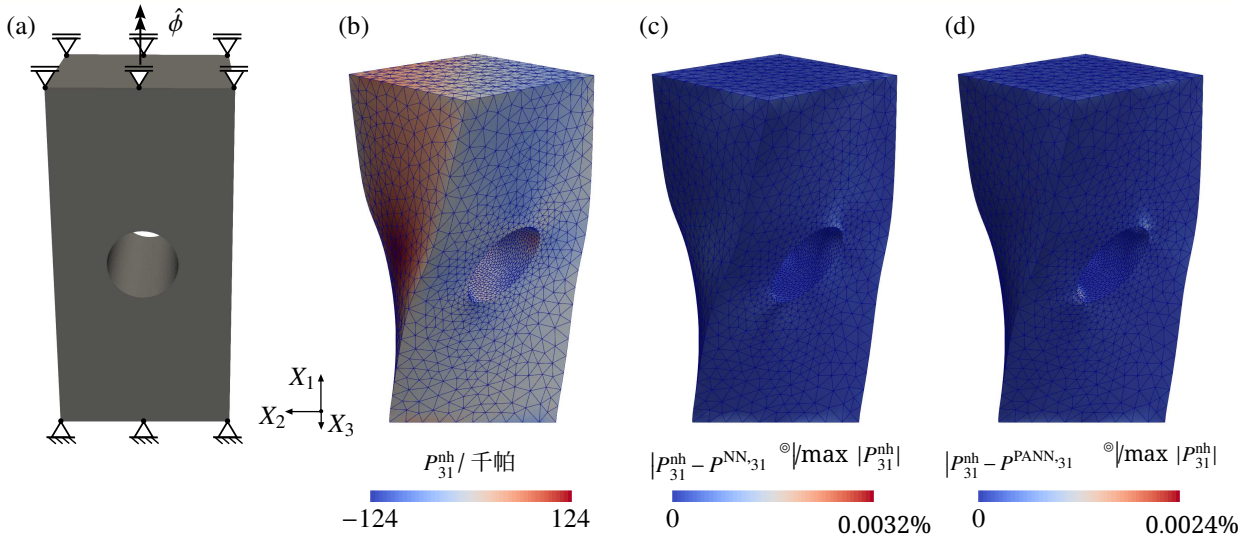


图 9：无翘曲扭转试样的有限元仿真：(a) 施加的边界条件，(b) 通过指定扭转角 $\hat{\phi} = 45°$ 得到的变形构形 B 上的宏观应力场 $P_{31}^{\mathrm{nh}}$，以及 (c) PANN 应力场 $P^{\mathrm{PANN},\circledcirc}_{,31}$ 的相对误差；(d) 基本条件模型的应力场 $P^{\mathrm{NN},\circledcirc}_{,31}$ 相对于 $P_{31}^{\mathrm{nh}}$ 的相对误差。神经网络使用多轴应力-应变数据 $\mathrm{D}^{\mathrm{red},\circledcirc}$、$|\mathrm{D}^{\mathrm{red},\circledcirc}| = 963$ 进行训练，并且每个本构方程仅采用一个隐藏层，其中包含 $N^{\mathrm{NN}} = 8$ 个神经元。

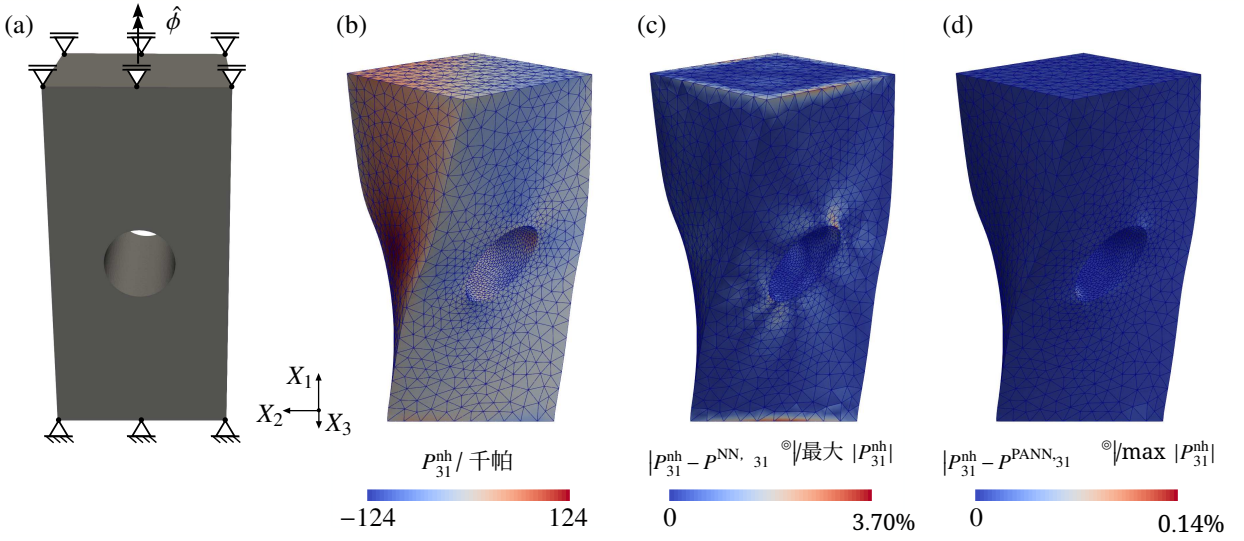


图10：无翘曲扭转试样的有限元仿真：(a) 施加的边界条件；(b) 通过指定 $\hat{\phi} = 45°$ 为扭转角，在变形构形 B 上得到的宏观应力场 $P_{31}^{\mathrm{nh}}$；(c) PANN 应力场的相对误差 $P^{\mathrm{PANN},\circledcirc}_{,31}$；以及 (d) 基本条件模型的应力场相对于 $P_{31}^{\mathrm{nh}}$ 的相对误差 $P^{\mathrm{NN},\circledcirc}_{,31}$。神经网络仅使用单轴应力-应变数据 $\mathrm{D}^{\mathrm{uniaxial},\circledcirc}$, $|\mathrm{D}^{\mathrm{uniaxial},\circledcirc}| = 15$进行训练，并且每个模型均以仅含一个隐藏层、具有 $N^{\mathrm{NN}} = 4$ 个神经元的结构实现为本构方程。

切线

$$\mathbb{C}^{\mathrm{NN},\circledcirc} := 4\frac{\partial^2 \psi^{\mathrm{NN},\circledcirc}}{\partial \boldsymbol{C} \partial \boldsymbol{C}} \in \mathcal{L}_4 \quad \text{and} \quad \mathbb{C}^{\mathrm{PANN},\circledcirc} := 4\frac{\partial^2 \psi^{\mathrm{PANN},\circledcirc}}{\partial \boldsymbol{C} \partial \boldsymbol{C}} \in \mathcal{L}_4 \,, \tag{66}$$

这些量是在通过标准牛顿-拉夫森方案求解时所必需的，可借助自动微分计算。

为评估神经网络的预测质量，以下将其与局部应力场 $P_{31}^{\mathrm{nh}}$ 进行比较，参见图 9(b)。我们首先考察由数据集 $\mathrm{D}^{\mathrm{red},\circledcirc}$ 中存储的复杂多轴应力-应变状态训练的模型，其在不变量空间中覆盖的定义域如图 8 所示。对于两个神经网络，相对误差均低于 0。004%，见图 9(c)、(d) 中的曲面图；也就是说，相对于参考应力 $P_{31}^{\mathrm{nh}}$，预测几乎完美。这并不令人意外，因为将扭转试样的变形映射到不变量空间后可以发现，该空间几乎完全被 $\mathrm{D}^{\mathrm{red},\circledcirc}$ 覆盖。因此，此处几乎不需要外推，另见 [22]中的讨论。接下来，我们仅考虑由数据集 $\mathrm{D}^{\mathrm{uniaxial},\circledcirc}$ 中存储的单轴应力状态训练的基于神经网络的模型，其在不变量空间中覆盖的曲线同样见图 8。观察图 10(c) 和 (d) 中的结果可知，基本条件模型的误差高达 3。7%，这是由于需要进行外推所致。不过，尽管如此，PANN模型仍实现了低于 0。14% 的极低误差。

因此，在本例中，基本条件模型和 PANN方法都能够很好地描述有限元仿真中较为复杂载荷工况下所学习到的非线性本构行为，而 PANN 表现更佳，尤其是在需要外推时。两者分别采用仅含一个隐藏层、具有$N^{\mathrm{NN}} = 8$ 和 $N^{\mathrm{NN}} = 4$ 个神经元的神经网络，与计算力学相关问题中通常使用的神经网络相比，这种规模非常小。此外，所实现的基本条件模型和 PANN 均能实现牛顿迭代典型的二次收敛，因此计算效率很高。

# 5 结论

本文提出了一种基于神经网络的可压缩有限应变超弹性本构模型。该方法称为 PANN，能够精确满足超弹性范畴内常见的所有本构条件，即热力学一致性、应力张量对称性、客观性、材料对称性、多凸性、增长条件，以及能量和应力归一化。此外，通过对物理上容许的变形状态空间进行采样，对神经网络势的非负性进行了数值检验。对于各向同性模型，基于解析考量缩小了所需的采样空间。所提出的模型能够描述高度非线性的超弹性关系，同时考虑其底层物理规律，并可使用 TensorFlow 等标准机器学习库进行训练。此外，本文还给出了可压缩 Neo-Hooke 势非负性的证明，据作者所知，此前尚未有人完成这一证明。

首先简要回顾基于神经网络的弹性模型，然后介绍有限应变超弹性，包括一般要求以及两类特定的各向异性类别及其模型。在此基础上，逐步构建 PANN 方法：将不同的不变量集作为凸神经网络的输入，使模型满足角动量平衡、客观性和材料对称性条件，同时满足热力学一致性和多凸性。随后，通过使用解析增长项满足体积增长条件。最后，通过多凸归一化项满足能量和应力归一化。应力归一化项取决于材料对称群，并以各向同性和横观各向同性材料行为为例进行推导。不过，满足物理条件的流程（例如使用归一化项）也可应用于解析方法或其他机器学习方法。之后，将展示

PANN 模型的适用性，并对利用解析势生成的各向同性和横观各向同性数据进行标定。所有情况下，即使面对高度多轴的变形状态以及含噪的应力-应变数据，也展现出极高的预测精度和稳健性。此外，结果表明，PANN 具有极强的外推能力。最后，还展示了其在有限元仿真中的直接应用。

总而言之，所介绍的用于可压缩有限应变超弹性的 PANN 模型已展现出高效工具的优势，可应用于固体力学领域的众多应用。因此，将物理学融入基于神经网络的模型至关重要，原因有几个：首先，这能带来可靠、即符合物理规律的模型预测。更重要的是，这对于提升神经网络的泛化性能也不可或缺 [24, 29] ，并能实现外推 [21, 27]。此外，只有纳入本构条件所带来的鲜明数学结构，才能利用工程应用中通常可获得的少量数据对模型进行标定。最后，这也使得采用规模相对较小的网络架构成为可能，参见 [21, 27]。因此，在构建基于神经网络的本构模型时，应尽可能精确地满足尽量多的本构条件，因为只有这样才能绝对确保这些条件得到满足。然而，对于某些应用，这可能无法实现，例如，在所考虑的对称群没有可用的不变量完备函数基，且采用不变量会过度限制模型灵活性时，参见 [26]。只有在这种情况下，才应通过近似满足某些本构条件来弱化模型结构，以获得更大的模型灵活性。除此之外，正如前文所述，精确满足本构条件所带来的结构性和可靠性应始终优先考虑。

为了推广我们提出的基于神经网络的弹性材料框架，未来计划进行若干扩展。例如，还需推导适用于更多材料对称群的多凸归一化项 [9]。此外，为了实现对潜在各向异性类型和方向的自动识别，采用张量基神经网络将是有益的补充 [13]。PANN模型在利用实验数据识别材料模型方面也很有前景 [11,57]。最后，还可以将其扩展至多物理场问题 [23, 27, 65] ，以拓展可能的应用领域。

## 致谢

Dominik K. Klein 和 Oliver Weeger 感谢德国研究联合会（DFG – 德国研究联合会）提供资助（项目编号：492770117），并感谢达姆施塔特工业大学计算工程中心下属计算工程研究生院给予的支持。

本文所展示的所有计算均在德累斯顿工业大学信息服务与高性能计算中心（ZIH）的 PC集群上完成。因此，作者感谢 ZIH 慷慨提供计算资源。最后，作者感谢 Franz Hirsch 和 Philipp Metsch 提供用于与高性能计算集群通信的服务器作业脚本。

## CRediT 作者贡献声明

**Lennart Linden：** 概念构思、正式分析、调查研究、方法论、可视化、软件、验证、可视化、撰写初稿、审阅与编辑。**Dominik Klein：** 概念构思、正式分析、方法论、可视化、软件、验证、撰写初稿、审阅与编辑。**Karl A. Kalina：** 概念构思、正式分析、方法论、可视化、软件、撰写初稿、审阅与编辑。**Jörg Brummund：** 正式分析，

方法论、审阅与编辑。**Oliver Weeger：** 概念构思、经费获取、资源、审阅与编辑。
**MarkusKästner：** 经费获取、资源、审阅与编辑。

## 声明

**利益冲突：**作者声明不存在任何利益冲突。

## 多层神经网络

在这项工作中，将不变量集合作为前馈神经网络 [1, 28] 的输入，其输出为标量值，并用于对超弹性势进行建模。第 3 节介绍了一种仅含一个隐藏层的前馈神经网络架构，事实证明，该架构足够灵活，能够满足许多实际应用的需求 [21, 27]。不过，第 3 节介绍的方法并不局限于仅含一个隐藏层的网络架构；为求完整起见，现介绍多层网络架构（参见图 11）。

简而言之，前馈神经网络可视为多个向量值函数的复合，其中各分量称为节点或神经元，而作用于每个节点的函数称为激活函数。前馈神经网络可通过两种方式提升灵活性：一是增加每个隐藏层中的节点数，如第 3 节对单隐藏层所做的那样；二是增加隐藏层的数量。将公式 (25) 中的单层架构推广到含有 $H$ 个隐藏层、且每个隐藏层中有$N^{\mathrm{NN},\,H}$ 个节点的网络架构，可得

$$\boldsymbol{A}_\alpha^{[1]} = \mathscr{F}\Big(\sum_{\beta=1}^{m} w_{\alpha\beta}^{[1]} I_\beta + \sum_{\gamma=1}^{A} w_{\alpha\gamma}^{*[1]} I_\gamma^* + b_\alpha^{[1]}\Big) \in \mathbb{R}^{N^{\mathrm{NN},1}}\,, \tag{67}$$

$$\boldsymbol{A}_\alpha^{[h]} = \mathscr{F}\Big(\sum_{\beta=1}^{N^{\mathrm{NN},h-1}} w_{\alpha\beta}^{[h]} \boldsymbol{A}_\beta^{[h-1]} + b_\alpha^{[h]}\Big) \in \mathbb{R}^{N^{\mathrm{NN},h}} \text{ with } h = 2,\dots,H\,, \tag{68}$$

$$\psi^{\mathrm{NN},\square}(\boldsymbol{\mathcal{I}}^*) = \sum_{\alpha=1}^{N^{\mathrm{NN},H}} W_\alpha \boldsymbol{A}_\alpha^{[H]} \in \mathbb{R}\,, \tag{69}$$

具有多凸、不可约且独立的不变量 $I_\beta(\boldsymbol{C},\ \mathrm{S}^{\square})$，以及第 3 节中定义的附加不变量$I_\gamma^*(\mathrm{I})$。通过在每个隐藏层中将激活函数 F 特别设定为凸函数且单调不减，并且 $W_\alpha \geq 0$，不变量的多凸性得以保持。同样，本文采用 *Softplus* 激活函数 $\mathscr{SP}(x) := \log(1+\exp(x))$；对于非负权重和任意偏置值，该函数是凸的且单调不减的，因此总体条件

$$W_\alpha, w_{\alpha\beta}^{[h]}, w_{\alpha\gamma}^{*[1]} \in \mathbb{R}_{\geq 0}, b_\alpha \in \mathbb{R} \quad \forall h \in \mathbb{N}_{\leq H}, \alpha \in \mathbb{N}_{\leq N^{\mathrm{NN},h-1}}, \beta \in \mathbb{N}_{\leq m}, \gamma \in \mathbb{N}_{\leq A} \tag{70}$$

最终得到一个多凸神经网络，参见注释 3.1；另请参阅 [26]，其中有更详尽的讨论和明确的证明。

## B 不变量的导数

为方便读者，表 4 给出了各向同性和横观各向同性不变量，以及它们关于 $\boldsymbol{C}$ 的导数。由此可得，行列式 $J = \sqrt{I_3}$ 的导数为

$$\frac{\partial J}{\partial \boldsymbol{C}} = \frac{J}{2}\boldsymbol{C}^{-1}\,, \tag{71}$$

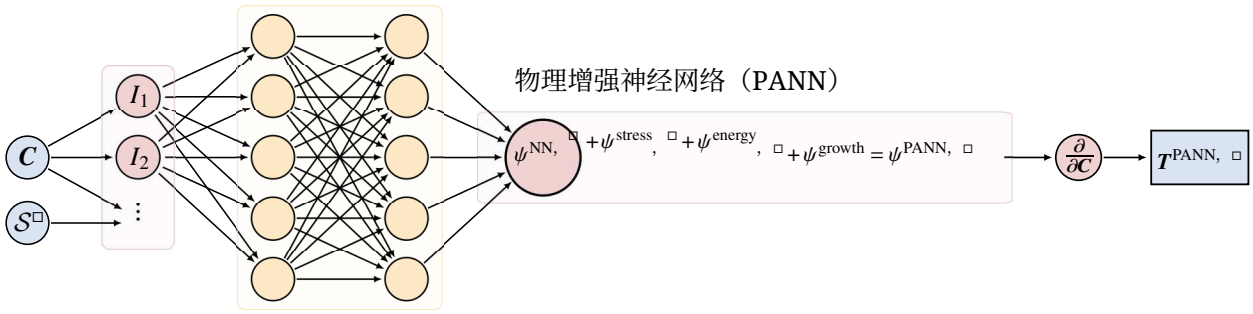


图11：所考虑的材料对称群 $\square$ 的多层 PANN 本构模型示意图。

表4：各向同性不变量和横观各向同性不变量的导数。

| | $I_\alpha$ | $\frac{\partial I_\alpha}{\partial \boldsymbol{C}}$ |
|---|---|---|
| 各向同性 | $I_1 := \mathrm{tr}\,\boldsymbol{C}$<br>$I_2 := \mathrm{tr}(\mathrm{cof}\,\boldsymbol{C})$<br>$I_3 := \det \boldsymbol{C}$ | $\mathbf{1}$<br>$I_1\mathbf{1}-\boldsymbol{C}$<br>$\mathrm{cof}\ \boldsymbol{C}$ |
| 横向<br>各向同性 | $I_4 := \mathrm{tr}(\boldsymbol{C}\cdot\boldsymbol{G})$<br>$I_5 := \mathrm{tr}(\mathrm{cof}(\boldsymbol{C})\cdot\boldsymbol{G})$ | $\boldsymbol{G}$<br>$I_5\boldsymbol{C}^{-1}-\mathrm{cof}(\boldsymbol{C})\cdot\boldsymbol{G}\cdot\boldsymbol{C}^{-1}$ |

这意味着，适配不变量 $I_1^* := -2J$ 的导数为

$$\frac{\partial I_1^*}{\partial \boldsymbol{C}} = -J\boldsymbol{C}^{-1}. \tag{72}$$

# C 应变能密度的非负性

在本附加章节中，我们通过解析考察讨论各向同性超弹性能函数的非负性。许多情况下，应变能密度函数以不变量表示，但需要注意的是，不变量空间中的任意状态并非都在物理上可接受，因为它们可能导致实数负值甚至复数的主伸长。因此，接下来有必要确定不变量的限制条件，以确保右柯西-格林变形张量正定。随后可以证明，对于某些应变能密度函数，它们仅会在物理上容许的变形状态下取非负值。下文将展示这一过程。

令 $I_1 := \mathrm{tr}\,\boldsymbol{C}$，$I_2 := \mathrm{tr}\,(\mathrm{cof}\ \boldsymbol{C})$ 和 $I_3 := \det \boldsymbol{C}$ 为右柯西-格林变形张量 $\boldsymbol{C} := \boldsymbol{F}^T\cdot\boldsymbol{F} \in \mathcal{Sym}$ 的主不变量。我们将右柯西-格林变形张量的特征值记为 $\kappa_1$，$\kappa_2$ 和 $\kappa_3$。这些特征值可由主伸长 $\lambda_1$，$\lambda_2$ 和 $\lambda_3$ 表示，它们是变形梯度 $\boldsymbol{F} \in \mathcal{GL}^+(3)$ 的特征值，即对所有 $\alpha \in \mathbb{N}_{\leq 3}$，$\kappa_\alpha = \lambda_\alpha^2$ 均成立。

**定义 C.1**。 如果变形张量 $\boldsymbol{C} \in \mathcal{Sym}$ 是正定的，即所有对应的特征值均为正（$\kappa_1,\ \kappa_2,\ \kappa_3 > 0$），则称该变形状态 为物理上可接受 。

**引理 C.2**。 特征方程

$$\kappa^3 - I_1\kappa^2 + I_2\kappa - I_3 = 0\,, \tag{73}$$

右侧的柯西　　一格林变形张量 $\boldsymbol{C} \in \mathcal{Sym}$ 当且仅当限制条件成立时，才具有实数特征值　　n

$$\Gamma := \left(\frac{q}{2}\right)^2 + \left(\frac{p}{3}\right)^3 \le 0\,, \quad \text{with } p := I_2 - \frac{I_1^2}{3}\,,\ q := -\frac{2I_1^3}{27} + \frac{I_1 I_2}{3} - I_3 \tag{74}$$

。

证明。 根据卡尔丹公式，当且仅当限制条件 (74) 成立时，三次方程 (73) 有实数解 [7]。□

**定理 C.3**。 当且仅当以下限制条件成立时，变形状态才是物理上可接受的

$$\Gamma(I_1, I_2, I_3) = \frac{1}{108}\left(4I_1^3 I_3 - I_1^2 I_2^2 + 4I_2^3 + 27I_3^2 - 18 I_1 I_2 I_3\right) \le 0 \quad \text{and} \quad I_1, I_2, I_3 > 0 \tag{75}$$

。

证明。”⇐”：已知限制条件 (75) 得到满足，且不变量 $I_1$、$I_2$、$I_3 > 0$ 均为正，因此可以通过反证法证明，特征值 $\kappa_\alpha$ 是方程 (73) 的正实数解。不变量分别为 $I_1 = \kappa_1 + \kappa_2 + \kappa_3$、$I_2 = \kappa_1\kappa_2 + \kappa_1\kappa_3 + \kappa_2\kappa_3$ 和 $I_3 = \kappa_1\kappa_2\kappa_3$，并假设方程 (73) 得到满足。根据引理 C.2，特征值为实数，因此不变量 $I_3$ 非负，当且仅当没有负特征值或恰有两个负特征值。进一步地，条件 $I_3 \neq 0$ 等价于对所有 $\alpha \in \mathbb{N}_{\le 3}$，均有 $\kappa_\alpha \neq 0$。因此，只需考察两个特征值为负、一个为正的情形。不失一般性，令 $\kappa_2$ 和 $\kappa_3$ 为负，$\kappa_1$ 为正（$\kappa_1 > 0$、$\kappa_2 < 0$、$\kappa_3 < 0$）。于是得到

$$I_2 = \kappa_1(\kappa_2 + \kappa_3) + \kappa_2\kappa_3 > 0\,, \tag{76}$$

$$\Leftrightarrow \frac{\kappa_2\kappa_3}{-(\kappa_2 + \kappa_3)} > \kappa_1 > -(\kappa_2 + \kappa_3)\,, \tag{77}$$

其中，最后一个不等式中用到了关系 $I_1 = \kappa_1 + \kappa_2 + \kappa_3 > 0$。乘以项 $-(\kappa_2 + \kappa_3) > 0$ 后，便得到矛盾

$$0 > \kappa_2^2 + \kappa_3^2 + \kappa_2\kappa_3\,, \tag{78}$$

因为右侧仅由正项组成。

”⇒”：这可直接由定义 $I_1 = \kappa_1 + \kappa_2 + \kappa_3$、$I_2 = \kappa_1\kappa_2 + \kappa_1\kappa_3 + \kappa_2\kappa_3$ 和 $I_3 = \kappa_1\kappa_2\kappa_3$ 以及卡尔丹公式 [7]得出。□

**引理 C.4**。 设条件 (75) 得到满足。当且仅当情形 $\Gamma = 0$ 成立时，至少有两个特征值 $\kappa_\alpha$ 相等。单特征值记为 $x > 0$，双重特征值记为 $y > 0$。因此，不变量为

$$I_1 = x + 2y > 0\,, \quad I_2 = 2xy + y^2 > 0\,, \quad I_3 = xy^2 > 0\,. \tag{79}$$

那么，偏导数为

$$\frac{\partial I_1}{\partial x} = 1 > 0\,, \qquad \frac{\partial I_1}{\partial y} = 2 > 0\,, \tag{80}$$

$$\frac{\partial I_2}{\partial x} = 2y > 0\,, \qquad \frac{\partial I_2}{\partial y} = 2x + 2y > 0\,, \tag{81}$$

$$\frac{\partial I_3}{\partial x} = y^2 > 0\,, \qquad \frac{\partial I_3}{\partial y} = 2xy > 0\,. \tag{82}$$

证明。第一个结果可直接由卡尔丹公式 [7] 得出。接着，偏导数可通过基本计算求得。□

## C.1 Neo-Hooke模型

**定理 C.5**。 应变能密度函数

$$\psi^{\mathrm{nh}}(I_1,I_3)=\frac{1}{2}\left(\mu\,(I_1-\ln I_3-3)+\frac{\lambda}{2}\,(I_3-\ln I_3-1)\right),\quad \mu,\lambda>0 \tag{83}$$

各向同性 的 *Neo-Hooke*模型 对于所有物理上容许的变形状态均非负 。

证明。连续可微函数 $\psi^{\mathrm{nh}}:\mathbb{R}^3\to\mathbb{R},(I_1,I_2,I_3)\mapsto\psi^{\mathrm{nh}}(I_1,I_3)$ 关于主不变量 $I_1,I_2,I_3$ 是凸的，且没有局部极值，因为关于主不变量的梯度不为零。注意，这个 Neo-Hooke模型不显式依赖主不变量 $I_2$ 。同样，如果将函数 $\psi^{\mathrm{nh}}(I_1,I_3)$ 的定义域限制为子集$\Omega:=\{(I_1,I_2,I_3)\in\mathbb{R}^3\mid\Gamma(I_1,I_2,I_3)\le 0,\ I_1,I_2,I_3>0\}$，则根据定理 C.3，梯度也不为零。因此，最小值位于容许域 $\Omega$ 的边界上，该容许域由下式给出

$$\partial\Omega_\Gamma:=\{(I_1,I_2,I_3)\in\mathbb{R}^3\mid\Gamma(I_1,I_2,I_3)=0,\ I_1>0,I_2>0,I_3>0\}\,, \tag{84}$$

$$\partial\Omega_1:=\{(I_1,I_2,I_3)\in\mathbb{R}^3\mid\Gamma(I_1,I_2,I_3)\le 0,\ I_1=0,I_2>0,I_3>0\}\,, \tag{85}$$

$$\partial\Omega_2:=\{(I_1,I_2,I_3)\in\mathbb{R}^3\mid\Gamma(I_1,I_2,I_3)\le 0,\ I_1>0,I_2=0,I_3>0\}\,, \tag{86}$$

$$\partial\Omega_3:=\{(I_1,I_2,I_3)\in\mathbb{R}^3\mid\Gamma(I_1,I_2,I_3)\le 0,\ I_1>0,I_2>0,I_3=0\}\,. \tag{87}$$

首先，我们注意到，对于 $I_1\to 0^+$ 或 $I_2\to 0^+$， $I_3\to 0^+$ 也成立，因为主不变量均为正实数（$I_1$， $I_2$， $I_3>0$）。根据公式（83），项 –ln $I_3$ 占主导，因此 $\partial\Omega_{1,\ 2,\ 3}$ 会导致 $\lim_{I_3\to 0^+}\psi^{\mathrm{nh}}(I_1,\ I_3)=\infty$，这不是局部极小值。

其次，我们考虑 $\partial\Omega_\Gamma$。因此，根据引理 C.4，我们现在只考虑独立变量 $x$， $y\in\mathbb{R}_{>0}$。为了求出 $\psi^{\mathrm{nh}}$ 在 $\partial\Omega_\Gamma$上的极值，我们计算势函数关于独立变量的偏导数，并令其为零:

$$\frac{\partial\psi^{\mathrm{nh}}}{\partial x}=\frac{\partial\psi^{\mathrm{nh}}}{\partial I_1}\frac{\partial I_1}{\partial x}+\frac{\partial\psi^{\mathrm{nh}}}{\partial I_3}\frac{\partial I_3}{\partial x}=\frac{\mu}{2}\left(1-\frac{1}{x}\right)+\frac{\lambda}{4}\left(1-\frac{1}{xy^2}\right)y^2=0\,, \tag{88}$$

$$\frac{\partial\psi^{\mathrm{nh}}}{\partial y}=\frac{\partial\psi^{\mathrm{nh}}}{\partial I_1}\frac{\partial I_1}{\partial y}+\frac{\partial\psi^{\mathrm{nh}}}{\partial I_3}\frac{\partial I_3}{\partial y}=\mu\left(1-\frac{1}{y}\right)+\frac{\lambda}{2}\left(1-\frac{1}{xy^2}\right)xy=0\,, \tag{89}$$

其中用到了 $x$， $y\neq 0$ 以及公式（80）–（82）。将公式（89）的一半从公式（88）中减去，得到方程

$$\frac{1}{2}(y-x)\left(-\frac{\mu}{xy}+\frac{\lambda}{2}\left(y-\frac{1}{xy}\right)\right)=0\,, \tag{90}$$

这显然对 $x=y$ 成立。另一个解可由下式得到

$$-\mu+\frac{\lambda}{2}\left(xy^2-1\right)=0\quad\Leftrightarrow\quad y^2=\frac{1}{x}\left(\frac{2\mu}{\lambda}+1\right), \tag{91}$$

但将其代入公式（88）后，会推出矛盾 $\mu=0$。

最后，我们将解 $x=y$ 代入公式（88），得到三次方程

$$x^3+\frac{2\mu}{\lambda}x-\left(\frac{2\mu}{\lambda}+1\right)=0\,, \tag{92}$$

这显然具有　　解 $x = 1$。利用多项式除法，我们可以得到二次方程

$$x^2 + x\left(\frac{2\mu}{\lambda} + 1\right) = 0\,, \tag{93}$$

这不会导出实数解，因为 $\mu,\ \lambda > 0$ 成立。事实上，对于体积变形状态 $\boldsymbol{C} = x\mathbf{1}$ 的三重特征值 $x \in (0,\ \infty)$，函数 $\psi^{\mathrm{nh}}(x)$ 是凸函数，因为对于球形张量的所有正三重特征值，$\frac{\partial^2 \psi^{\mathrm{nh}}}{\partial x^2} > 0$ 成立。

因此，$x = y = 1$ 是 $\psi^{\mathrm{nh}}(x,\ y)$ 的全局最小点，因而也是在容许集合 $\Omega$ 上方程（88）和（89）的唯一实数且非负解。最后，可以确定 $\psi^{\mathrm{nh}}(x = 1,\ y = 1) = 0$，即未变形状态具有零能量。因此，对于所有物理上容许的变形状态，$\psi^{\mathrm{nh}}(I_1,\ I_3) \geq 0$ 成立。□

### C.2 各向同性 PANN

**定理 C.6**。考虑各向同性 *PANN* 模型

$$\psi^{\mathrm{PANN},\circledcirc}(\boldsymbol{\mathcal{I}}^*) = \sum_{\alpha=1}^{N^{\mathrm{NN}}} W_\alpha\, \mathcal{SP}\left(\sum_{\beta=1}^{3} w_{\alpha\beta} I_\beta + w^*_{\alpha 1} I^*_1 + b_\alpha\right) + \left(J + \frac{1}{J} - 2\right)^2 - \mathfrak{n}(J-1) + \psi^{\mathrm{energy},\circledcirc}\,, \tag{94}$$

包括归一化常数

$$\mathfrak{n} := 2\left(\frac{\partial \psi^{\mathrm{NN},\circledcirc}}{\partial I_1} + 2\frac{\partial \psi^{\mathrm{NN},\circledcirc}}{\partial I_2} + \frac{\partial \psi^{\mathrm{NN},\circledcirc}}{\partial I_3} + \frac{\partial \psi^{\mathrm{NN},\circledcirc}}{\partial I^*_1}\frac{\partial I^*_1}{\partial I_3}\right)\Bigg|_{\boldsymbol{C}=\mathbf{1}} \in \mathbb{R}\,, \tag{95}$$

其中 $\mathrm{I}^* := (I_1,\ I_2,\ I_3,\ I^*_1)$ 及附加不变量 $I^*_1 := -2\sqrt{I_3} = -2J$。已知

$$\frac{\partial \psi^{\mathrm{PANN},\circledcirc}}{\partial I_1} > 0 \quad \text{and} \quad \frac{\partial \psi^{\mathrm{PANN},\circledcirc}}{\partial I_2} > 0 \tag{96}$$

成立时，势函数 $\psi^{\mathrm{PANN},\circledcirc}(\mathrm{I}^*)$ 仅在满足 $\lambda \in (0,\ \infty)$ 的体积变形状态 $\boldsymbol{C} = \lambda^2\mathbf{1}$ 下才可能存在局部极小值。

证明。连续可微函数 $\psi^{\mathrm{PANN},\circledcirc}\colon \mathbb{R}^3 \to \mathbb{R},\ (I_1, I_2, I_3) \mapsto \psi^{\mathrm{PANN},\circledcirc}(I_1, I_2, I_3, -2\sqrt{I_3})$ 关于主不变量 $I_1, I_2, I_3$ 是凸的，且没有局部极值，因为根据条件 (96)，关于这些不变量的梯度不会消失。需要指出的是，假设 (96) 对神经网络的近似质量并无太大限制，因为至少对一个 $\alpha, \beta \in \mathbb{N}_{\leq N^{\mathrm{NN}}}$，较小的正权重 $w_{\alpha 1}, W_\alpha$ 和 $w_{\beta 2}, W_\beta$ 已满足这些假设。因此，当函数 $\psi^{\mathrm{PANN},\circledcirc}(\mathrm{I}^*)$ 的定义域限制为容许不变量子集 $\Omega := \{(I_1, I_2, I_3) \in \mathbb{R}^3 \mid \Gamma(I_1, I_2, I_3) \leq 0,\ I_1, I_2, I_3 > 0\}$ 时，梯度不会消失，如定理 C.3 所述。

由于增长项 (32) 已包含在势函数 $\psi^{\mathrm{PANN},\circledcirc}(\mathrm{I}^*)$ 中，且所有其他项均为有限值，因此无需考虑方程 (85)–(87) 中的边界 $\partial\Omega_1, \partial\Omega_2, \partial\Omega_3$。论证过程与定理 C.5 的证明类似。因此，只需考虑方程 (84) 中给出的边界 $\partial\Omega_\Gamma$。多凸神经网络模型关于不变量的偏导数记为

$$\psi_\gamma := \frac{\partial \psi^{\mathrm{PANN},\circledcirc}}{\partial I_\gamma}\,. \tag{97}$$

为求局部极值，将借助引理 C.4 表示的关于特征值的偏导数设为零：

$$\frac{\partial \psi^{\mathrm{PANN},\circledcirc}}{\partial x} = \psi_1 + 2y\,\psi_2 + y^2\,\psi_3 = 0\,, \tag{98}$$

$$\frac{\partial \psi^{\mathrm{PANN},\circledcirc}}{\partial y} = 2\,\psi_1 + 2(x+y)\,\psi_2 + 2xy\,\psi_3 = 0\,. \tag{99}$$

将式 (99) 的一半从式 (98) 中减去，得到

$$(y-x)\,(\psi_2 + y\,\psi_3) = 0\,. \tag{100}$$

将第一个解 $\psi_2 = -y\,\psi_3$ 代入式 (98)，会得到矛盾 $\psi_1 = y^2\,\psi_3$ 和 $\psi_2 = -y\,\psi_3$；这也无法满足 $y > 0$ 以及假设 $\psi_1 > 0$ 和 $\psi_2 > 0$。因此，另一个解 $x = y$ 是式 (98) 和式 (99) 的唯一解。

与上文的 Neo-Hooke模型 相同，对 $\psi^{\mathrm{PANN}}$、$^{\circledcirc}(\mathrm{I}^{*})$ 的全局最小值的搜索可以缩小到球形张量 $\boldsymbol{C} = x\mathbf{1}$ 的正实三重特征值 $x \in (0,\infty)$ 所构成的（凸）集合。特别地，按构造，$x = 1$ 是一个局部最小点，且 $\psi^{\mathrm{PANN}}$、$^{\circledcirc}(x = 1) = 0$，因为式 (37) 中给出的归一化项的归一化常数 $\mathfrak{n}$ 是据此选取的。然而，由于 $J = x^{3/2}$ 在 $x \in (0,\infty)$ 上严格凹，并且因归一化可能带有正或负因子，因此通常无法证明 $\psi^{\mathrm{PANN}}$、$^{\circledcirc}(x)$ 关于三重特征值是凸的。因此，最小点 $x = 1$ 可能不唯一，并且球形张量 $\boldsymbol{C} = x\mathbf{1}$ 还可能存在其他三重特征值 $x \neq 1$，使得导数 $\frac{\partial \psi^{\mathrm{PANN}}}{\partial x}^{\circledcirc}$ 为零，且 $\psi^{\mathrm{PANN}}$、$^{\circledcirc}(x) < 0$ 成立。□

# D 随机

本附录部分展示了第 4.2.2 节所述统计研究中获得的原始数据。图 12(a)–(b) 分别给出了各向同性（□ = ◎）和横观各向同性（□ = k）情形的直方图，其中包含中位数、第25百分位数和第75百分位数。对于本统计研究中的每个模型，共完成了 300 轮训练。基于缩减数据集 $\mathrm{D}^{\mathrm{red},\,\square}$，每轮训练中分别对相应的神经网络进行 30 次训练，并保存具有最低均方误差（57）的最优训练状态的参数，如第 3.4 节所述。

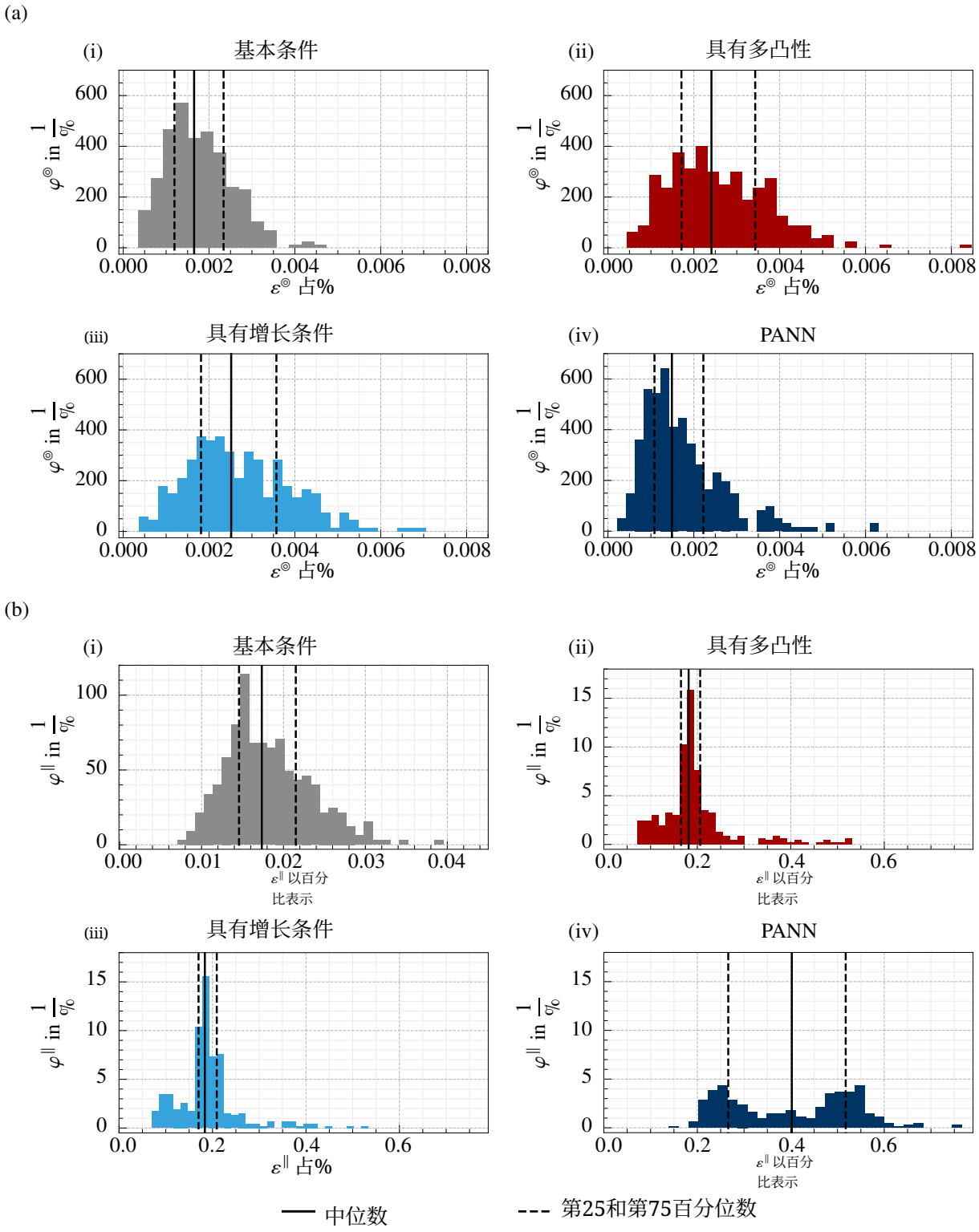


图12：不同各向异性下，方程（64）中误差度量 $\varepsilon^{\square}$ 的 $N^{\text{bins}} := 30$ 直方图：(a) 各向同性和 (b) 基于不变量的横观各向同性模型。(i) 基本条件，(ii) 多凸性，(iii) 增长条件和多凸性，以及 (iv) 满足包括归一化在内的所有条件的 PANN。结果通过 300 轮训练生成，每轮从 30 次训练中选取最佳结果。使用缩减数据集 $\mathrm{D}^{\text{red}, \square}$进行了标定。

# 参考文献

[1] C. C. Aggarwal. 神经网络与深度学习. 第1版。Springer International Publishing，2018。[2] M. Alnæs 等。“FEniCS 项目 1.5 版”。载于：数值软件档案 3.100（2015）。[3] B. Amos、L. Xu 和 J. Z. Kolter。“输入凸神经网络”。载于：第*34*届国际机器学习会议论文集。D. Precup 和 Y. W. Teh 编。第70卷。机器学习研究论文集。PMLR，2017，第146–155页。arXiv：1609.07152。[4] F. As’ad、P. Avery 和 C. Farhat。“一种用于数据驱动本构建模的力学信息驱动人工神经网络方法”。载于：*International Journal for Numerical Methods in Engineering* 123.12（2022），第2738–2759页。数字对象标识符：10.1002/nme.6957。[5] J. M. Ball。“非线性弹性静力学中的本构不等式与存在性定理”。载于：赫瑞瓦特专题讨论会：非线性分析与力学。R.J. Knops 编。第1卷。伦敦：Pitman，1977，第187–241页。[6] J. M. Ball。“非线性弹性力学中的凸性条件与存在性定理”。载于：*Archive for Rational Mechanics and Analysis*63.4（1976），第337–403页。数字对象标识符：10.1007/BF00279992。[7] Ilja N. Bronshtein、Konstantin A. Semendyayev、Gerhard Musiol 和 Heiner Muehlig。数学手册。第5版。柏林：施普林格，2007。[8] P. Chen 和 J. Guilleminot。“用于超弹性本构模型的多凸神经网络：一种修正方法”。载于：《力学研究通讯》 125（2022），第103993页。数字对象标识符：10.1016/j.mechrescom.2022.103993。[9] V. Ebbing。“面向所有各向异性类别的多凸能量函数设计”。博士论文。杜伊斯堡-埃森大学，2010。[10] M. Fernández、M. Jamshidian、T. Böhlke、K. Kersting 和 O. Weeger。“结合材料理论与数据驱动方法的有限变形各向异性超弹性本构模型及其在立方晶格超材料中的应用”。载于：《计算力学》 67.2 （ 2021 ），653–677。数字对象标识符：10.1007/s00466-020-01954-7。[11] M. Flaschel 、 S. Kumar 和 L. De Lorenzis。“可解释超弹性本构定律的无监督发现”。载于：*Computer Methods in Applied Mechanics and Engineering* 381（2021），第113852页。数字对象标识符：10.1016/j.cma.2021.113852。[12] A. L. Frankel、R. E. Jones 和 L. P. Swiler。“超弹性材料的张量基高斯过程模型”。载于：《建模与计算机器学习期刊》 1.1（2020）。数字对象标识符：10.1615/.2020033325。[13] J. N. Fuhg、N. Bouklas 和 R. E. Jones。“通过张量基神经网络从数据中学习超弹性各向异性”。载于：*Journal of the Mechanics and Physics of Solids* 168（2022），第105022页。数字对象标识符：10.1016/j.jmps.2022.105022。[14] J. N. Fuhg、M. Marino 和 N. Bouklas。“数据驱动本构模型的局部近似高斯过程回归：开发及与神经网络的比较”。载于：*Computer Methods in Applied Mechanics and Engineering* 388（2022），第114217页。数字对象标识符：10.1016/j.cma.2021.114217。[15] C. Geuzaine 和 J.-F. Remacle。“Gmsh：一款内置前处理与后处理功能的三维有限元网格生成器”。载于：*International Journal for Numerical Methods in Engineering* 79.11（2009），第1309–1331页。数字对象标识符：https://doi.org/10.10

02/nme.2579.

[16] J. Ghaboussi、J. H. Garrett 和 X. Wu。“基于神经网络的材料行为知识建模”。载于：《工程力学杂志》117.1 (1991)，第 132–153 页。数字对象标识符：10.1061/(ASCE)0733-9399(1991)117:1(132).[17] T. Gärtner、M. Fernández 和 O. Weeger。“基于人工神经网络的各向异性均匀化本构模型弹性梁格构非线性多尺度模拟”。载于：《计算力学》68 (2021 年 11 月)。数字对象标识符：10.1007/s00466-021-02061-x.[18] P. Haupt。连续介质力学与材料理论。第 2 版。Springer Berlin Heidelberg，2002 年。[19] G. A. Holzapfel。非线性固体力学：工程连续介质方法。第 2 版。威利，2000 年。[20] S. Huang、Z. He、B. Chem 和 C. Reina。“变分奥恩萨格神经网络 (VONNs)：一种用于非平衡偏微分方程的基于热力学的变分学习策略”。载于：*Journal of the Mechanics and Physics of Solids* 163 (2022)，第 104856 页。数字对象标识符：10.1016/j.jmps.2022.104856。[21] K. A. Kalina、L. Linden、J. Brummund 和 M. Kästner。“FEANN：一种基于物理约束神经网络和自动化数据挖掘的高效数据驱动多尺度方法”。载于：《计算力学》(2023)。数字对象标识符：10.1007/s00466

-022-02260-0.

[22] K. A. Kalina、L. Linden、J. Brummund、P. Metsch 和 M. Kästner。“基于人工神经网络的各向同性超弹性自动化本构建模”。载于：《计算力学》69（2022年1月），第1–20页。数字对象标识符：10.1007/s00466-021-

02090-6.

[23] K. A. Kalina、P. Metsch、J. Brummund 和 M. Kästner。“基于微观模拟的磁流变弹性体宏观模型”。英文。发表于：《国际固体与结构杂志》193-194（2020），第 200–212 页。数字对象标识符：10.1016/j.ijsolstr.2020.02.028。

[24] G. E. Karniadakis、I. G. Kevrekidis、L. Lu、P. Perdikaris、S. Wang 和 L. Yang。“物理信息机器学习”。发表于：*NatureReviews Physics*（2021）。数字对象标识符：10.1038/s42254-021-00314-5。[25] T. Kirchdoerfer 和 M. Ortiz。“数据驱动计算力学”。发表于：*Computer Methods in Applied Mechanics and Engineering* 304（2016），第 81–101 页。数字对象标识符：10.1016/j.cma.2016.02.001。[26] D. K. Klein、M. Fernández、R. J. Martin、P. Neff 和 O. Weeger。“基于神经网络的多凸各向异性超弹性”。发表于：*Journal of theMechanics and Physics of Solids* 159（2022），第 104703 页。数字对象标识符：10.1016/j.jmps.2021.104703。[27] D. K. Klein、R. Ortigosa、J. Martínez-Frutos 和 O. Weeger。“ 采用物理增强神经网络的有限电弹性 ”。发表于：*Computer Methods in Applied Mechanics and Engineering* 400（2022），第 115501 页。[28] S. Kollmannsberger、D. D’Angella、M. Jokeit 和 L. Herrmann。计算力学中的深度学习。第 977 卷。Studies in Computational Intelligence。施普林格，2021。数字对象标识符：10.1007/978-3-030-76587-3。[29] S. Kumar 和 D. M. Kochmann。“机器学习能为计算固体力学做些什么”。发表于：计算力学的当前趋势与开放问题。F. Aldakheel、B. Hudobivnik、M. Soleimani、H. Wessels、C. Weißenfels 和 M. Marino 编。施普林格国际出版：Springer International Publishing，2022，第 275–285 页。数字对象标识符：10.1007/978-3-030-87312-7_27。[30] B. A. Le、J. Yvonnet 和 Q.-C. He。“利用神经网络对非线性弹性材料进行计算均质化”。发表于：*International Journal forNumerical Methods in Engineering* 104.12（2015），第 1061–1084 页。数字对象标识符：10.1002/nme.4953。[31] S. Lehmich、P. Neff 和 J. Lankeit。“关于函数 $C \mapsto f(\det C)$ 在正定矩阵上的凸性”。发表于：*Mathematics and Mechanics of Solids* 19.4（2014），第 369–375 页。数字对象标识符：10.1177/1081286512466099。[32] G. Liang 和 K. Chandrashekhara。“弹性体泡沫的神经网络本构模型”。发表于：*Engineering Structures* 30.7（2008），第 2002–2011 页。数字对象标识符：10.1016/j.engstruct.2007.12.021。[33] L. Linden、K. A. Kalina、J. Brummund、P. Metsch 和 M. Kästner。“基于人工神经网络的各向同性超弹性热力学一致本构建模”。发表于：*PAMM* 21.1（2021）。数字对象标识符：10.1002/pamm.202100144。[34] J. Ling、R. Jones 和 J. Templeton。“具有不变性属性的系统的机器学习策略”。发表于：*Journal of Computational Physics*318（2016），第 22–35 页。数字对象标识符：10.1016/j.jcp.2016.05.003。[35] K. Linka、M. Hillgärtner、K. Abdolazizi、R. Aydin、M. Itskov 和 C. Cyron。“本构人工神经网络：通过深度学习实现预测性数据驱动本构建模的一种快速通用方法”。发表于：*Journal of Computational Physics*（2020），第 110010 页。数字对象标识符：10.1016/j.jcp.2020.110010。[36] K. Linka 和 E. Kuhl。“迈向自动化模型发现的新型本构人工神经网络”。发表于：*Computer Methods in Applied Mechanics and Engineering* 403（2023），第 115731 页。数字对象标识符：10.1016/j.cma.2022.115731。[37] Z. Liu、Y. Du、Y. Chen 和 M. Tegmark。“物理增强学习：超越物理信息学习的新范式”。发表于：2021。[38] A. Logg、K.-A. Mardal 和 G. Wells。利用有限元法自动求解微分方程：*FEniCS* 实用指南。第 84 卷。Springer Science & Business Media，2012。[39] R. J. Martin、I.-D. Ghiba 和 P. Neff。“对数型亨基应变能的多凸延拓”。发表于：*Analysis and Applications* 17.03（2019），第 349–361 页。数字对象标识符：10.1142/S0219530518500173。[40] F. Masi 和 I. Stefanou。“采用基于热力学的人工神经网络（TANN）对非弹性材料进行多尺度建模”。发表于：*Computer Methods in Applied Mechanics andEngineering* 398（2022），第 115190 页。数字对象标识符：10.1016/j.cma.2022.1

15190。

[41] J. Merodio 和 P. Neff。“关于纤维增强非线性弹性固体的拉伸失稳与椭圆性丧失”。载于：力学档案 58（2006），第 293–303 页。[42] C. Naumann 和 J. Ihlemann。“关于重现穆林斯效应的伪弹性材料模型的热力学”。载于：《国际固体与结构杂志》（2015）。数字对象标识符：10.1016/j.ijsolstr.2015.05.014。[43] P. Neff、I.-D. Ghiba 和 J. Lankeit。“指数化亨基对数应变能。第一部分：本构问题与秩一凸性”。载于：*Journal of Elasticity*121（2015），第 143–234 页。数字对象标识符：10.1007/s10659-015-9524-7。[44] J. Nocedal 和 S. J. Wright。“序列二次规划”。载于：《数值优化》。施普林格纽约，2006，第 529–562 页。数字对象标识符：10.1007/978-0-387-40065-5_18。[45] R. W. Ogden、G. Saccomandi 和 I. Sgura。“将超弹性模型拟合到实验数据”。载于：《计算力学》34.6（2004），第 484–502 页。数字对象标识符：10.1007/s00466-004-0593-y。[46] G. C. Y. Peng 等。“多尺度建模与机器学习相遇：我们能学到什么？”载于：《工程计算方法档案》（2020）。数字对象标识符：10.1007/s11831-020-09405-5。[47] A. Ricker 和 P. Wriggers。“弹性体超弹性连续体模型的系统拟合与比较”。载于：《工程计算方法档案》（2023）。数字对象标识符：10.1007/s11831-022-09865-x。

[48] K. Sagiyama 和 K. Garikipati。“机器学习材料物理学：利用来自马氏体微观结构的弹性自由能数据训练的深度神经网络，能够高精度预测均质化应力场”。载于：（2019）。arXiv：1901.00524 [cond-mat, physics:physics]。[49] J. Schröder 和 P. Neff。“基于多凸自由能函数的超弹性横向各向同性不变量表述”。载于：《国际固体与结构杂志》40（2003），第 401–445 页。数字对象标识符：10.1016/S0020-7683(02)00458-4。[50] J. Schröder、P. Neff 和 V. Ebbing。“基于晶体学动机结构张量的各向异性多凸能量”。载于：*Journal of the Mechanics andPhysics of Solids* 56.12（2008），第 3486–3506 页。数字对象标识符：https://doi.org/10.1016/j.jmp

s.2008.08.008

[51] Y. Shen、K. Chandrashekhara、W. F. Breig 和 L. R. Oliver。“用于橡胶材料的基于神经网络的本构模型”。载于：《橡胶化学与技术》77.2 (2004)，第 257–277 页。数字对象标识符：10.5254/1.3547822。[52] B. Smith、F. De Goes 和 T. Kim。“稳定的 Neo-Hookean 软组织模拟”。载于：《*ACM*图形学汇刊》37.2 (2018)，第 1–15 页。数字对象标识符：10.1145/3180491。[53] S. R. St. Pierre、K. Linka 和 E. Kuhl。“基于主伸长的本构神经网络自主发现了人脑组织 Ogden 模型的一个子类”。载于：《脑多物理学》4 (2023)，第 100066 页。数字对象标识符：10.1016/j.brain.2023.100066。[54] V. Tac、K. Linka、F. Sahli-Costabal、E. Kuhl 和 A. B. Tepole。“物理信息驱动的数据驱动超弹性基准测试”。载于：(2023)。数字对象标识符：10.48550/ARXIV.2301.10714。[55] V. Tac、F. Sahli Costabal 和 A. B. Tepole。“采用多凸神经常微分方程的数据驱动组织力学”。载于：*Computer Methods inApplied Mechanics and Engineering* 398 (2022)，第 115248 页。数字对象标识符：10.1016/j.cma.2022.115248

.

[56] V. Tac、V. D. Sree、M. K. Rausch 和 A. B. Tepole。“各向异性软生物组织力学行为的数据驱动建模”。载于：*Engineering with Computers* (2022)。数字对象标识符：10.1007/s00366-022-01733-3。[57] P. Thakolkaran、A. Joshi、Y. Zheng、M. Flaschel、L. De Lorenzis 和 S. Kumar。“NN-EUCLID：无需应力数据的深度学习超弹性”。载于：*Journal of the Mechanics and Physics of Solids* 169 (2022)，第 105076 页。数字对象标识符：10.1016/j.jmps.2022.105076。[58] C. Truesdell 和 W. Noll。《非线性力学场论》。第 3 版。Springer Berlin Heidelberg，2004。[59] N. N. Vlassis、R. Ma 和 W. C. Sun。“计算力学中的几何深度学习（第一部分）：各向异性超弹性”。载于：*Computer Methods in Applied Mechanics and Engineering*371 (2020)，第 113299 页。数字对象标识符：10.1016/j.cma.2020.113299。[60] N. N. Vlassis 和 W. C. Sun。“用于可解释弹塑性模型（含水平集硬化）的热力学信息神经网络 Sobolev 训练”。载于：*Computer Methods in Applied Mechanics and Engineering* 377 (2021)，第 113695 页。数字对象标识符：10.1016/j.cma.2021.113695。[61] N. N. Vlassis、P. Zhao、R. Ma、T. Sewell 和 W. C. Sun。“基于分子动力学推断的迁移学习模型：用于单斜晶体有限应变超弹性的 Sobolev 训练及物理约束验证”。载于：*International Journal for Numerical Methods in Engineering* 123.17 (2022)，第 3922–3949 页。数字对象标识符：10.1002/nme.6992。[62] P. Weber、J. Geiger 和 W. Wagner。“受约束神经网络训练及其在超弹性材料建模中的应用”。载于：《计算力学》68.5 (2021)，第 1179–1204 页。数字对象标识符：10.1007/s00466-021-02064-8。[63] D. Yang Gao、P. Neff、I. Roventa 和 C. Thiel。“ 关于右柯西 - 格林张量中非线性弹性能的凸性 ”。载于：*Journal of Elasticity*127 (2017)，第 303–308 页。数字对象标识符：10.1007/s10659-016-9601-6。[64] L. Zee 和 E. R. Sternberg。“不可压缩超弹性固体平衡理论中的常规椭圆性与强椭圆性”。载于：*Archive for Rational Mechanics and Analysis* 83 (1983)，第 53–90 页。数字对象标识符：10.1007/BF00281087。[65] M. Zlatić 和 M. Čanađija。“不可压缩橡胶热弹性：一种神经网络方法”。载于：《计算力学》71.5 (2023)，第 895–916 页。数字对象标识符：10.1007/s00466-023-02278-y。[66] C. Zopf 和 M. Kaliske。“通过基于神经网络的方法对未硫化弹性体进行数值表征”。载于：*Computers & Structures* 182 (2017)，第 504–525 页。数字对象标识符：10.1016/j.compstruc.2016.12.012。