通过西格玛点权重的循环元适应实现鲁棒无迹卡尔曼滤波

Kenan Majewski*、Michał Modzelewski、Marcin Zugaj、
Piotr Lichota航空与应用力学研究所 华沙理工大学波兰华沙
*kenan.majewski.dokt@pw.edu.pl

arXiv:2603.04360v1 [cs.LG] 4 Mar 2026

摘要—**无迹卡尔曼滤波器（UKF）是非线性状态估计中广泛使用的工具；然而，其性能受限于无迹变换（UT）的静态参数化。由固定缩放参数决定的传统加权方案假设系统具有隐含的高斯性，无法适应时变动力学或重尾测量噪声。本文提出元自适应无迹卡尔曼滤波器（MA-UKF），这是一种将西格玛点权重合成重新构建为超参数优化问题，并通过带记忆的元学习加以解决的框架。与依赖瞬时启发式校正的标准自适应滤波器不同，我们的方法采用循环上下文编码器，将测量新息的历史压缩为紧凑的潜在嵌入。该嵌入为策略网络提供信息，使其能够在每个时间步动态合成西格玛点的均值与协方差权重，从而有效控制滤波器对预测与测量的信任程度。通过滤波器的递归逻辑对系统进行端到端优化，MA-UKF 学会在保持估计一致性的同时最大限度地提高跟踪精度。对机动目标进行的数值基准测试表明，MA-UKF 显著优于标准基线，对非高斯闪烁噪声表现出更强的鲁棒性，并能有效泛化至训练期间未见过的分布外动态工况。**

索引词—**无迹卡尔曼滤波器、元学习、自适应滤波、非线性估计、鲁棒跟踪**

I. 引言

自 Julier 和 Uhlmann 提出以来，无迹卡尔曼滤波器（UKF）已成为非线性状态估计的标准方法 [1]。通过将一组确定性的西格玛点传播通过非线性系统动力学，UKF 能够针对高斯输入捕获三阶精度的后验矩，从而避免扩展卡尔曼滤波器（EKF）固有的线性化误差，而无需进行显式的雅可比计算 [2]。然而，标准 UKF 的性能严格受无迹变换（UT）参数化方式的制约，具体而言，取决于决定西格玛点分布范围和权重的缩放参数（α、 β、 κ）。尽管早期研究已探索基于模型的学习方法来优化这些西格玛点 [3],，但传统实现仍主要依赖预先选定的静态参数先验，其依据是对高斯性和稳态噪声统计特性的隐含假设。因此，滤波器的适应性在以机动目标、非平稳动力学或重尾测量噪声为特征的复杂环境中受到严重限制 [4]。

为克服这些局限，自适应卡尔曼滤波领域历来依赖启发式噪声估计方法，例如 Sage-Husa估计器 [5] 或交互多模型（IMM）架构 [6]。尽管交互多模型通过在不同动态模型之间切换来增强鲁棒性，但其计算开销较大，且依赖预先定义的有限模式集。近年来，信号处理与深度学习的融合催生了更广泛的人工智能辅助卡尔曼滤波器范式 [7]。过去，递归估计器与神经网络之间的关系往往是反向的，例如使用扩展卡尔曼滤波器显式训练循环神经网络（RNN）的权重 [8]。如今，深度学习正积极增强状态估计逻辑。近期文献涉及用于机动目标的基于模型的深度学习 [9],、用于跟踪的自监督深度状态空间模型 [10], , 以及将循环神经网络与卡尔曼滤波器变体融合的混合架构，以实现复杂轨迹预测 [11] 和实时传感器漂移补偿 [12]。

与此同时，可微滤波的发展使滤波器参数能够通过反向传播进行端到端优化 [13], [14]。这一理念已成功拓展至多种贝叶斯估计器，包括用于可微粒子滤波器的端到端学习 [15] ，以及用于点质量滤波器的神经网络增强型自适应网格 [16]。对于连续状态空间，KalmanNet 等方法 [17] 能够从数据流中学习动态调整卡尔曼增益或噪声协方差。尽管这些方法提升了状态估计性能，但许多方法基本上仍将底层的西格玛点几何视为固定的归纳偏置。

认识到这一不足后，近期研究开始明确致力于无迹变换的自适应。近期提出的框架已将数据驱动的未知输入估计应用于西格玛点滤波器 [18], , 并利用循环神经网络动态自动调整标准 UKF 的超参数 [19]。此外，近期研究还提出利用神经网络微调高斯滤波器中采样点的显式空间位置 [20]。然而，这些方法仍主要侧重于调整标准标量约束，或依赖瞬时启发式校正，尚未将无迹变换的整体构造重新表述为连续的超控制

本研究已提交给 IEEE，供其考虑发表。版权可能在不另行通知的情况下转让，届时此版本可能不再可供访问。

问题。

借助基于梯度的元学习领域的进展 [21], [22], 尤其是将“学会学习”视为一种通过记忆增强架构实现的时间序列建模问题的奠基性研究 [23], [24],我们提出了一种新框架，将西格玛点权重的调节重新表述为序贯决策过程。在此上下文中，滤波器被视为可微计算图，其超参数会针对未见过的动态工况进行实时调整。这标志着从静态的基于模型的估计转向上下文感知元学习，在这种方法中，适应逻辑被整合到循环神经网络的隐藏状态中。

在本文中，我们提出了元自适应无迹卡尔曼滤波器（MA-UKF）。与绕过状态空间建模的端到端神经网络架构不同，我们的框架在保持贝叶斯递推结构完整性的同时，通过学习策略自适应地调节无迹变换参数。我们架构的核心是一个循环上下文编码器 [25],，它将测量新息的历史压缩为潜在嵌入。该嵌入使系统无需显式的模式切换逻辑，即可区分动态机动与传感器异常。随后，策略网络将这一上下文映射为西格玛点的最优均值与协方差权重，同时满足无迹变换的数学凸性约束。通过使用时间反向传播解析梯度进行端到端训练 [26], ，MA-UKF 学会了在即时跟踪精度与长期估计稳定性之间取得平衡。

本文的贡献包括以下三个方面：

1) **可微元滤波：**我们将无迹变换的参数化表述为可微计算图中的双层优化问题。这使得能够通过时间反向传播解析梯度，端到端学习最优的数据驱动西格玛点权重。

2) **记忆增强自适应：**我们引入循环上下文编码器，将新息历史压缩为潜在嵌入。该时间上下文驱动学习策略实时动态调节西格玛点权重，从而有效区分真实机动与传感器异常。

3) **鲁棒性与分布外泛化：**我们证明，在重尾闪烁噪声下，MA-UKF显著优于经过优化的标准滤波器和交互多模型基线。至关重要的是，学习策略展现出强大的分布外泛化能力，能够对训练期间未见的高敏捷机动保持准确跟踪。

本文其余部分安排如下：第二节介绍理论框架，详细阐述无迹变换和元学习目标。第三节介绍所提出的 MA-UKF 架构和训练方法。第四节展示实验设置和对比结果。最后，第五节总结全文并展望未来研究方向。

II. 理论框架

A. 问题定义

考虑一个离散时间非线性动态系统的状态估计问题。标准滤波方法假设系统服从典型的高斯统计特性，而实际应用中往往会遇到工况变化和分布不匹配。设真实生成过程定义如下：

$$\mathbf{x}_k = f(\mathbf{x}_{k-1}) + \mathbf{w}_k, \quad \mathbf{w}_k \sim \mathcal{D}_w(\cdot), \tag{1}$$

$$\mathbf{z}_k = h(\mathbf{x}_k) + \mathbf{v}_k, \quad \mathbf{v}_k \sim \mathcal{D}_v(\cdot), \tag{2}$$

其中，$\mathbf{x}_k \in \mathbb{R}^{n_x}$ 是时间步 k处的状态向量，$\mathbf{z}_k \in \mathbb{R}^{n_z}$是量测向量，而 $f: \mathbb{R}^{n_x} \to \mathbb{R}^{n_x}$ 和 $h: \mathbb{R}^{n_x} \to \mathbb{R}^{n_z}$ 分别为非线性转移函数和观测函数。

在标准表述中，假设过程噪声 $\mathbf{w}_k$ 和测量噪声 $\mathbf{v}_k$ 服从固定的高斯分布，即 $\mathcal{D}_w = \mathcal{N}(\mathbf{0}, \mathbf{Q})$ 和 $\mathcal{D}_v = \mathcal{N}(\mathbf{0}, \mathbf{R})$。然而，在鲁棒跟踪场景中（例如机动目标或雷达角闪烁），真实分布 $\mathcal{D}_w$ 和 $\mathcal{D}_v$ 会随时间变化、呈多峰特征，或具有非零偏度和高峰度。

我们的目标是推导一个由 θ参数化的最优估计器 $\mathcal{F}_\theta$，使其在每个时间步都能生成状态估计 $\hat{\mathbf{x}}_k$。我们旨在最小化有限时域 T内的期望累积估计误差：

$$\theta^* = \arg\min_\theta \mathbb{E}_{\tau \sim \mathcal{P}_\tau} \left[\sum_{k=1}^{T} \|\mathbf{x}_k - \mathcal{F}_\theta(\hat{\mathbf{x}}_{k-1}, \mathbf{z}_k)\|_2^2 \right], \tag{3}$$

其中，$\tau = (\mathbf{x}_{0:T}, \mathbf{z}_{1:T})$表示从真实系统分布 $\mathcal{P}_\tau$中采样得到的轨迹。不同于传统的自适应滤波（其试图显式估计时变噪声协方差 $\mathbf{Q}$ 或 $\mathbf{R}$），我们寻求通过参数集 θ直接优化滤波器自身的内部积分参数。

B. 标准无迹卡尔曼滤波器公式

无迹卡尔曼滤波器使用一组$2n_x + 1$ 西格玛点 $\mathcal{X}_{k-1} = \{\mathcal{X}_{k-1}^i\}_{i=0}^{2n_x}$来近似状态分布，并为均值和协方差分别配以权重 $W_i^{(m)}$ 和 $W_i^{(c)}$。给定上一步的后验估计 $\hat{\mathbf{x}}_{k-1}$ 和协方差 $\mathbf{P}_{k-1}$，西格玛点的生成方式如下：

$$\mathcal{X}_{k-1}^0 = \hat{\mathbf{x}}_{k-1} \tag{4}$$

$$\mathcal{X}_{k-1}^i = \hat{\mathbf{x}}_{k-1} + \left[\sqrt{(n_x + \lambda)\mathbf{P}_{k-1}}\right]_i, \quad i = 1, \ldots, n_x \tag{5}$$

$$\mathcal{X}_{k-1}^{i+n_x} = \hat{\mathbf{x}}_{k-1} - \left[\sqrt{(n_x + \lambda)\mathbf{P}_{k-1}}\right]_i, \quad i = 1, \ldots, n_x \tag{6}$$

其中 $[\cdot]_i$ 表示矩阵平方根的第 i列，通常通过下三角乔列斯基分解计算。尺度参数定义为 $\lambda = \alpha^2(n_x + \kappa) - n_x$，其中 α 控制西格玛点的散布范围， κ 为次级尺度参数。

这些点对应的常数权重由下式给出:

$$W_0^{(m)} = \frac{\lambda}{n_x + \lambda}, \tag{7}$$

$$W_0^{(c)} = \frac{\lambda}{n_x + \lambda} + (1 - \alpha^2 + \beta), \tag{8}$$

$$W_i^{(m)} = W_i^{(c)} = \frac{1}{2(n_x + \lambda)}, \quad i = 1 \dots 2n_x. \tag{9}$$

其中 β 是用于融入底层分布先验知识的标量参数（对于最优高斯近似，通常为 $\beta = 2$ ）。

预测步骤： 西格玛点通过非线性过程函数进行传播:

$$\mathcal{X}_{k|k-1}^i = f(\mathcal{X}_{k-1}^i). \tag{10}$$

通过加权重组计算先验 状态均值和协方差:

$$\hat{\mathbf{x}}_k^- = \sum_{i=0}^{2n_x} W_i^{(m)} \mathcal{X}_{k|k-1}^i, \tag{11}$$

$$\mathbf{P}_k^- = \sum_{i=0}^{2n_x} W_i^{(c)} (\mathcal{X}_{k|k-1}^i - \hat{\mathbf{x}}_k^-)(\mathcal{X}_{k|k-1}^i - \hat{\mathbf{x}}_k^-)^\top + \mathbf{Q}_k. \tag{12}$$

更新步骤： 随后将传播后的西格玛点代入测量函数:

$$\mathcal{Z}_{k|k-1}^i = h(\mathcal{X}_{k|k-1}^i). \tag{13}$$

预测测量均值 $\hat{\mathbf{z}}_k^-$ 、新息协方差$\mathbf{P}_{zz}$和互协方差$\mathbf{P}_{xz}$ 计算如下:

$$\hat{\mathbf{z}}_k^- = \sum_{i=0}^{2n_x} W_i^{(m)} \mathcal{Z}_{k|k-1}^i, \tag{14}$$

$$\mathbf{P}_{zz} = \sum_{i=0}^{2n_x} W_i^{(c)} (\mathcal{Z}_{k|k-1}^i - \hat{\mathbf{z}}_k^-)(\mathcal{Z}_{k|k-1}^i - \hat{\mathbf{z}}_k^-)^\top + \mathbf{R}_k, \tag{15}$$

$$\mathbf{P}_{xz} = \sum_{i=0}^{2n_x} W_i^{(c)} (\mathcal{X}_{k|k-1}^i - \hat{\mathbf{x}}_k^-)(\mathcal{Z}_{k|k-1}^i - \hat{\mathbf{z}}_k^-)^\top. \tag{16}$$

最后，计算卡尔曼增益$\mathbf{K}_k$ ，并更新后验估计:

$$\mathbf{K}_k = \mathbf{P}_{xz} \mathbf{P}_{zz}^{-1}, \tag{17}$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{z}_k - \hat{\mathbf{z}}_k^-), \tag{18}$$

$$\mathbf{P}_k = \mathbf{P}_k^- - \mathbf{K}_k \mathbf{P}_{zz} \mathbf{K}_k^\top. \tag{19}$$

C. 可微滤波范式

可微滤波体现了贝叶斯递归估计与深度学习的融合。传统上，卡尔曼滤波器和神经网络被视为彼此独立的实体:前者是基于模型的估计器，后者是数据驱动近似器。在可微滤波范式中，滤波器本身被表示为一个计算图，类似于循环神经网络，但其归纳偏置受到物理运动方程和贝叶斯更新规则的严格约束 [13], [14], [27]。

令 $\mathcal{F}_\theta$ 表示由方程 (4) 至 (19) 定义的单步无迹卡尔曼滤波器更新函数。我们可以将时刻 k 的滤波器状态视为一个元组 $\mathbf{s}_k = (\hat{\mathbf{x}}_k, \mathbf{P}_k)$。滤波器的演化可表示为一个可微递归:

$$\mathbf{s}_k = \mathcal{F}_\theta(\mathbf{s}_{k-1}, \mathbf{z}_k), \tag{20}$$

其中， θ 表示可调参数（例如控制西格玛点缩放的策略网络权重）。重要的是，对于正定矩阵，计算式（5）和（6）中矩阵平方根所需的乔列斯基分解等操作均可微。这使梯度能够连续地流经滤波器的随机协方差参数 [28]。

为优化 θ，我们计算累积损失 $\mathcal{L}$ 相对于参数的梯度。由于参数θ 会影响整个轨迹，因此我们采用时间反向传播（BPTT） [29]。应用链式法则后，总梯度是各时间步梯度的累积:

$$\frac{d\mathcal{L}}{d\theta} = \sum_{k=1}^{T} \frac{\partial \mathcal{L}_k}{\partial \mathbf{s}_k} \frac{d\mathbf{s}_k}{d\theta}, \tag{21}$$

其中，总状态导数 $\frac{d\mathbf{s}_k}{d\theta}$ 通过递归计算:

$$\frac{d\mathbf{s}_k}{d\theta} = \frac{\partial \mathcal{F}_\theta}{\partial \theta} + \frac{\partial \mathcal{F}_\theta}{\partial \mathbf{s}_{k-1}} \cdot \frac{d\mathbf{s}_{k-1}}{d\theta}. \tag{22}$$

这里, $\frac{\partial \mathcal{F}_\theta}{\partial \mathbf{s}_{k-1}}$ 是无迹卡尔曼滤波器更新步骤的雅可比矩阵，精确反映不确定性 $\mathbf{P}_{k-1}$ 如何通过系统非线性 k传播。通过展开这一计算图，并利用自动微分优化 θ ，我们在保留卡尔曼滤波器结构可解释性的同时，获得了神经网络由数据驱动的适应能力。

D. 上下文驱动的元策略构式

标准 UKF 的一个根本局限在于 UT 的参数化是静态的。权重 $\mathcal{W} = \{W^{(m)}, W^{(c)}\}$使西格玛点的几何结构保持时间不变，隐含地假设系统的非线性和噪声统计特性保持平稳。在动态环境中，这种僵化迫使系统在跟踪敏捷性（高响应性）与噪声抑制（高稳定性）之间作出权衡。

为解决这一问题，我们在可微分编程 [24]的范式下重新构建滤波问题。我们不再将 UKF 视为固定算法，而是将其视为一个由元策略 π_θ参数化的有向计算图。该策略充当超控制器，根据估计上下文动态调节滤波器的积分参数。

在每个时间步 k，策略都会处理测量创新 $\boldsymbol{\nu}_k$。这一代理信号用于更新循环隐藏状态$\mathbf{h}_k$。这个潜在状态概括了轨迹的历史信息，使网络能够区分短暂机动与非物理传感器噪声。随后，合成自适应权重如下:

$$\mathbf{h}_k = \text{RNN}_\theta(\boldsymbol{\nu}_k, \mathbf{h}_{k-1}), \tag{23}$$

$$\mathcal{W}_k = \{W_k^{(m)}, W_k^{(c)}\} = \pi_\theta(\mathbf{h}_k). \tag{24}$$

这些动态权重会立即注入标准 UKF 的预测与更新周期，以生成后验状态估计 $\hat{\mathbf{x}}_k$。

该架构构成一个双层优化问题 [30]。内层循环 执行递归贝叶斯状态估计，而外层循环 则优化策略参数 θ。由于无迹卡尔曼滤波器中的组成运算——包括矩阵乘法、Cholesky 分解和线性方程组求解——几乎处处可微，因此可以通过时间反向传播解析计算累积估计误差 $\nabla_\theta \mathcal{L}$ 的梯度。

III. 元自适应无迹卡尔曼滤波器

本节详细介绍所提出的元自适应无迹卡尔曼滤波器（MA-UKF）架构。该框架旨在将物理状态传播与滤波器的统计参数化解耦。如图 1 所示，系统由三个功能模块组成：新息特征提取、循环上下文编码和凸权重合成。

A. 新息特征提取

在递归估计中，新息序列 $\boldsymbol{\nu}_k$ 是自适应的主要信号。经典自适应滤波器会利用归一化新息平方（NIS）等标量指标触发启发式逻辑 [31]，而原始新息幅值对传感器模态、目标距离和尺度高度敏感。

为确保数值稳定性并促进梯度的高效传播，我们采用可学习的特征提取阶段。为避免循环依赖，我们使用前一时间步的名义几何离散度计算代理新息 $\tilde{\boldsymbol{\nu}}_k \in \mathbb{R}^{n_z}$：

$$\tilde{\boldsymbol{\nu}}_k = \mathbf{z}_k - \sum_{i=0}^{2n_x} W_{i,k-1}^{(m)} \mathcal{Z}_{k|k-1}^{i}. \tag{25}$$

通过线性投影，随后进行层归一化 [32]，将这一原始信号投影到高维特征空间：

$$\mathbf{e}_k = \mathrm{ReLU}(\mathrm{LayerNorm}(\mathbf{W}_{\text{in}}\tilde{\boldsymbol{\nu}}_k + \mathbf{b}_{\text{in}})), \tag{26}$$

其中$\mathbf{W}_{\text{in}} \in \mathbb{R}^{d_h \times n_z}$ 是可学习的投影矩阵，$\mathbf{b}_{\text{in}} \in \mathbb{R}^{d_h}$ 是偏置，$\mathbf{e}_k \in \mathbb{R}^{d_h}$ 表示瞬时残差的静态特征向量。在闪烁噪声环境下，这种归一化对于确保鲁棒性至关重要，因为离群测量值可能比名义方差高出数个数量级。

B. 循环上下文编码

仅凭瞬时残差不足以区分测量离群值与机动的开始。尽管这两种现象都会产生较大的范数 $\|\boldsymbol{\nu}_k\|$，但它们呈现出不同的时间结构：闪烁通常是不相关且短暂的，而机动则表现为相关的低频趋势。

为捕捉这些时间依赖性，我们采用门控循环单元（GRU）。GRU 充当非线性无限冲激响应（IIR）滤波器，随时间整合信息以更新潜在上下文嵌入$\mathbf{h}_k \in \mathbb{R}^{d_h}$。其更新动力学由下式决定：

$$\mathbf{u}_k = \sigma_g(\mathbf{W}_u \mathbf{e}_k + \mathbf{U}_u \mathbf{h}_{k-1} + \mathbf{b}_u), \tag{27}$$

$$\mathbf{r}_k = \sigma_g(\mathbf{W}_r \mathbf{e}_k + \mathbf{U}_r \mathbf{h}_{k-1} + \mathbf{b}_r), \tag{28}$$

$$\tilde{\mathbf{h}}_k = \tanh(\mathbf{W}_h \mathbf{e}_k + \mathbf{U}_h(\mathbf{r}_k \odot \mathbf{h}_{k-1}) + \mathbf{b}_h), \tag{29}$$

$$\mathbf{h}_k = (1 - \mathbf{u}_k) \odot \mathbf{h}_{k-1} + \mathbf{u}_k \odot \tilde{\mathbf{h}}_k, \tag{30}$$

其中$\sigma_g(\cdot)$表示 Sigmoid 激活函数，$\odot$ 表示哈达玛积，$\mathbf{u}_k$ 和$\mathbf{r}_k$ 分别为更新门和重置门。参数矩阵定义如下：$\mathbf{W}_u, \mathbf{W}_r, \mathbf{W}_h \in \mathbb{R}^{d_h \times d_h}$ 为输入权重，$\mathbf{U}_u, \mathbf{U}_r, \mathbf{U}_h \in \mathbb{R}^{d_h \times d_h}$ 为循环权重；相应的偏置为$\mathbf{b}_u, \mathbf{b}_r, \mathbf{b}_h \in \mathbb{R}^{d_h}$。这一机制使策略能够有选择地抑制高频噪声，同时保留对真实动态偏移的敏感性。

C. 凸西格玛点权重合成

MA-UKF 的核心贡献是学习得到的策略映射 $\pi_\theta : \mathbf{h}_k \to \mathcal{W}_k$，它能够动态合成无迹变换参数。首先，将高维隐藏状态投影到低维上下文流形 $\mathbf{c}_k \in \mathbb{R}^{d_p}$：

$$\mathbf{c}_k = \mathrm{ReLU}(\mathrm{LayerNorm}(\mathbf{W}_{proj}\mathbf{h}_k + \mathbf{b}_{proj})), \tag{31}$$

其中 $\mathbf{W}_{proj} \in \mathbb{R}^{d_p \times d_h}$ 和 $\mathbf{b}_{proj} \in \mathbb{R}^{d_p}$。

无迹滤波中的一项关键约束是协方差更新的数值稳定性，这要求协方差矩阵保持半正定。标准参数化（例如使用负的 λ 值）可能产生非正定矩阵，导致数值不稳定，并可能在训练过程中引发乔列斯基分解失败。为保证稳定性，我们对西格玛点权重施加凸性约束 。均值权重 $W^{(m)}$ 和协方差权重 $W^{(c)}$ 通过 Softmax函数生成：

$$\mathbf{l}_k^{(m)} = \mathbf{W}_\pi^{(m)} \mathbf{c}_k + \mathbf{b}_\pi^{(m)}, \tag{32}$$

$$W_{i,k}^{(m)} = \frac{\exp(l_{i,k}^{(m)})}{\sum_{j=0}^{2n_x} \exp(l_{j,k}^{(m)})}, \tag{33}$$

其中 $\mathbf{W}_\pi^{(m)} \in \mathbb{R}^{(2n_x+1)\times d_p}$ 和$\mathbf{b}_\pi^{(m)} \in \mathbb{R}^{2n_x+1}$。协方差权重 $W^{(c)}$ 的计算方式类似。该约束确保 $\sum_i W_i = 1$ 和 $W_i > 0$，从而保证预测协方差 $\mathbf{P}_k^-$ 保持正定。虽然这一限制使滤波器无法捕捉负权重所能体现的高阶峰度效应，但它提供了端到端可微训练所需的严格稳定性保证，确保在整个优化过程中，无迹卡尔曼滤波器中的乔列斯基分解始终有效。

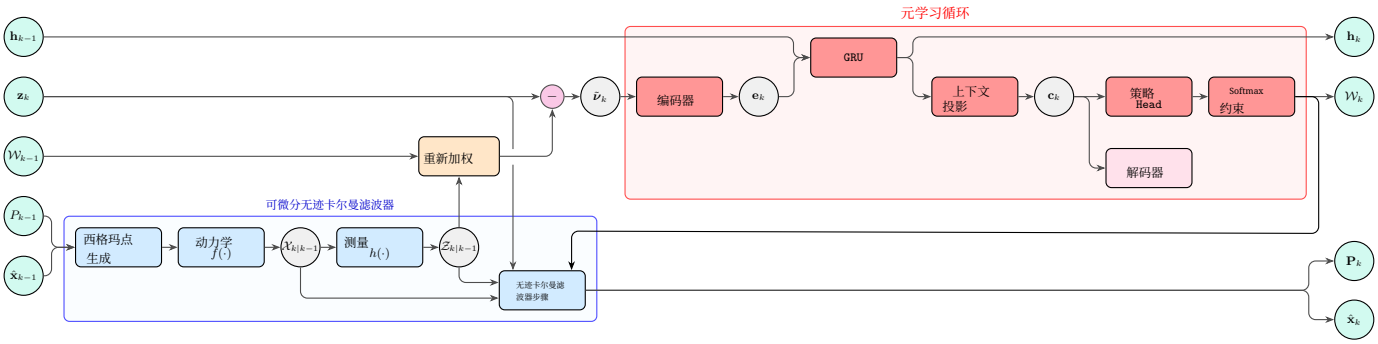

图1. Meta-Adapti的架构 ve无迹卡尔曼滤波器（MA-UKF）。循环上下文编码器处理新息历史，以生成潜在上下文 t
嵌入，该嵌入驱动一个策略网络 k，以实时调制西格玛点权重。

D. 算法与计算流程

MA-UKF 的递归执行过程总结于算法 1。其一项关键设计原则是将优化负担解耦：计算开销巨大的元策略训练在线下完成，从而可以通过轻量级循环神经网络的一次高效前向传播，在线执行滤波器参数的动态实时自适应。

Algorithm 1 Meta-Adaptive UKF Recursive Step

Require: $\hat{\mathbf{x}}_{k-1}, \mathbf{P}_{k-1}, \mathcal{W}_{k-1}, \mathbf{h}_{k-1}, \mathbf{z}_k, \mathbf{Q}_k, \mathbf{R}_k, \gamma$

$\mathcal{X}_{k-1} \leftarrow \left[\hat{\mathbf{x}}_{k-1}, \hat{\mathbf{x}}_{k-1} \pm \sqrt{\gamma \mathbf{P}_{k-1}}\right]$

$\mathcal{X}_{k|k-1} \leftarrow f(\mathcal{X}_{k-1})$

$\mathcal{Z}_{k|k-1} \leftarrow h(\mathcal{X}_{k|k-1})$

$\tilde{\boldsymbol{\nu}}_k \leftarrow \mathbf{z}_k - \sum_{i=0}^{2n_x} W_{i,k-1}^{(m)} \mathcal{Z}_{k|k-1}$

$\mathbf{e}_k \leftarrow \mathrm{ReLU}(\mathrm{LayerNorm}(\mathbf{W}_{in}\tilde{\boldsymbol{\nu}}_k + \mathbf{b}_{in}))$

$\mathbf{h}_k \leftarrow \mathrm{GRU}(\mathbf{e}_k, \mathbf{h}_{k-1})$

$\mathcal{W}_k = \{W_k^{(m)}, W_k^{(c)}\} \leftarrow \pi_\theta(\mathbf{h}_k)$

$\hat{\mathbf{x}}_k^- \leftarrow \sum_{i=0}^{2n_x} W_{i,k}^{(m)} \mathcal{X}_{k|k-1}^i$

$\mathbf{P}_k^- \leftarrow \sum_{i=0}^{2n_x} W_{i,k}^{(c)} (\mathcal{X}_{k|k-1}^i - \hat{\mathbf{x}}_k^-)(\mathcal{X}_{k|k-1}^i - \hat{\mathbf{x}}_k^-)^\top + \mathbf{Q}_k$

$\hat{\mathbf{z}}_k^- \leftarrow \sum_{i=0}^{2n_x} W_{i,k}^{(m)} \mathcal{Z}_{k|k-1}^i$

$\mathbf{P}_{zz} \leftarrow \sum_{i=0}^{2n_x} W_{i,k}^{(c)} (\mathcal{Z}_{k|k-1}^i - \hat{\mathbf{z}}_k^-)(\mathcal{Z}_{k|k-1}^i - \hat{\mathbf{z}}_k^-)^\top + \mathbf{R}_k$

$\mathbf{P}_{xz} \leftarrow \sum_{i=0}^{2n_x} W_{i,k}^{(c)} (\mathcal{X}_{k|k-1}^i - \hat{\mathbf{x}}_k^-)(\mathcal{Z}_{k|k-1}^i - \hat{\mathbf{z}}_k^-)^\top$

$\mathbf{K}_k \leftarrow \mathbf{P}_{xz}\mathbf{P}_{zz}^{-1}$

$\hat{\mathbf{x}}_k \leftarrow \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{z}_k - \hat{\mathbf{z}}_k^-)$

$\mathbf{P}_k \leftarrow \mathbf{P}_k^- - \mathbf{K}_k \mathbf{P}_{zz} \mathbf{K}_k^\top$

return $\hat{\mathbf{x}}_k, \mathbf{P}_k, \mathcal{W}_k, \mathbf{h}_k$

E. 计算复杂度分析

对于实时跟踪应用而言，计算效率至关重要。元自适应模块引入的额外开销由特征提取、GRU 状态更新和策略投影层决定，其带来的额外时间复杂度为 $\mathcal{O}(d_h^2 + d_h n_z + d_h d_p + n_x d_p)$。鉴于所提出架构的维度较低（$n_x = 5$，$n_z = 2$，$d_h = 32$，$d_p = 16$），这一神经推理序列在每个递归周期中严格只需数千次浮点运算（FLOPs）。借助高度优化的稠密矩阵运算，这一前向传播在当代微处理器上的延迟低于微秒级。因此，与标准单模型 UKF 相比，MA-UKF 的总执行时间带来的额外开销可以忽略不计；标准单模型 UKF 的性能瓶颈仍在于 $\mathcal{O}(n_x^3)$ 乔列斯基分解和协方差更新。至关重要的是，这使 MA-UKF 能够实现稳健的自适应，同时完全避免交互多模型架构评估 M 个并行假设时固有的 $\mathcal{O}(M \cdot n_x^3)$ 倍乘性扩展。

F. 通过端到端解析梯度进行训练

我们将策略优化表述为在由 N 条轨迹构成的数据集上的监督回归问题。损失函数 $\mathcal{L}_\theta$ 包含一个主要跟踪误差项和一个辅助正则化项：

$$\mathcal{L}_\theta = \sum_{k=1}^{T} \|\mathbf{x}_k - \mathcal{F}_\theta(\hat{\mathbf{x}}_{k-1}, \mathbf{z}_k)\|_2^2 + \lambda_{aux} \|\tilde{\boldsymbol{\nu}}_k - g(\mathbf{c}_k)\|_2^2. \quad (34)$$

这里，$g(\cdot)$ 是一个由 λ_{aux} 加权的辅助解码器，用于从潜在上下文$\mathbf{c}_k$重构新息。这种正则化促使潜在状态$\mathbf{h}_k$ 保留新息序列的丰富表示，从而防止训练过程中发生模式崩溃。

IV. 实验评估

为了验证所提出的 MA-UKF 的有效性，我们进行了一系列数值实验，重点考察存在重尾测量噪声和未建模动态机动时的鲁棒跟踪。我们将该方法与标称UKF、超参数优化无迹卡尔曼滤波器以及交互多模型滤波器进行基准比较。

A. 仿真环境和数据集

我们考虑一个二维雷达跟踪场景，其中状态向量 $\mathbf{x} = [p_x, v_x, p_y, v_y, \omega]^\top$ 按照协调转弯（CT）运动学模型演化，采样间隔为离散时间

为 $\Delta t = 0.1$ 秒。雷达提供带噪声的距离和方位角测量：

$$\mathbf{z}_k = \left[\sqrt{p_x^2 + p_y^2},\ \arctan(p_y/p_x)\right] + \mathbf{v}_k. \tag{35}$$

1) 训练方案（随机匀速转弯）：MA-UKF 在通过标准匀速转弯模型生成的随机轨迹合成数据集上进行训练。对于每条轨迹，初始状态 $\mathbf{x}_0$ 均随机生成：位置 p_x，p_y 从 $\mathcal{U}(-1000,\ 1000)$ m 中采样；速度则由速率 $v \sim \mathcal{U}(10,\ 30)$ m/s 和航向角 $\varphi \sim \mathcal{U}(0,\ 2\pi)$ rad 确定。转弯率从 $\omega \sim \mathcal{U}([-0.5, -0.1] \cup [0.1, 0.5])$ rad/s 中抽取，以确保模型针对明显的左转和右转机动进行训练，而非接近零的转弯率。噪声 $\mathbf{v}_k$ 从高斯混合模型（GMM）中抽取：

$$\mathbf{v}_k \sim (1-\epsilon)\mathcal{N}(0, \mathbf{R}) + \epsilon\mathcal{N}(0, \eta\mathbf{R}), \tag{36}$$

其中 $\epsilon = 0.1$ 表示闪烁离群值（信号尖峰）的概率为 10%，$\mathbf{R}$ 是标称传感器噪声协方差，$\eta = 20$ 是闪烁缩放因子。

2) 评估机制（分布外泛化）：尤为重要的是，评估在代表训练期间未见动态工况的数据集上进行。我们采用高敏捷蛇形机动，其中目标经历参数随机化的正弦加速度：

$$\mathbf{a}(t) = \left[A_x \sin(\omega_x t),\ A_y \cos(\omega_y t)\right]^\top. \tag{37}$$

对于每个评估序列，振幅 A_x、A_y 从 $\mathcal{U}([-20, -10] \cup [10, 20])$ m/s^2 中采样，频率 ω_x、ω_y 从 $\mathcal{U}(-2, 2)$ rad/s中采样。

该轨迹违反了滤波器状态转移函数（$f(\cdot)$）所固有的恒定转弯率假设。此外，我们在评估期间将测量离群值的严重程度加倍，并把闪烁缩放因子设为 $\eta = 40$，以测试滤波器的鲁棒性。未建模动力学与极端噪声的这一组合严格检验了学习策略的泛化能力。

B. 实现与基线

所有模型均使用Python，借助JAX和Equinox库实现可微分编程。MA-UKF在NVIDIA RTX 3060 GPU上使用Adam优化器（$lr = 10^{-3}$）训练了1800个训练轮次，批大小为64，序列长度为60。

基线1：标称UKF。我们将与针对标称性能调优的标准 UKF进行比较。虽然无迹变换通常采用较小的散布（$\alpha = 10^{-3}$）来近似局部线性化，但此类设置在高噪声条件下数值不稳定。因此，我们采用 [1]提出的参数化：$\alpha = 1.0$、$\beta = 2.0$和 $\kappa = 3 - n_{x\circ}$ 该设置将西格玛点置于距均值 $\pm\sqrt{3\sigma}$ 处，从而捕捉高斯分布的四阶矩，并提供足够的几何散布。过程噪声协方差$\mathbf{Q}$ 和测量噪声 $\mathbf{R}$ 固定为标称真实值。

基线2：优化的无迹卡尔曼滤波器（无迹卡尔曼滤波器*）。为建立严格的基线，我们利用 Optuna 框架在 100 次试验中进行超参数优化搜索。所得参数（$\alpha \approx 17.26$、$\beta \approx 2.59$、$\kappa \approx 0.15$）代表静态无迹卡尔曼滤波器在该特定数据集上性能的经验上界。

基线 3：交互多模型无迹卡尔曼滤波器。：我们还采用了一个交互多模型滤波器，将两种无迹卡尔曼滤波器模式结合起来：用于直线运动的恒定速度（CV）模式，以及用于机动的协调转弯模式 [33]。转移概率矩阵的对角线元素固定为 $\Pi_{ii} = 0.95$。我们还评估了经过超参数优化的版本（IMM-UKF*）。

所提出的 MA-UKF。策略网络采用隐藏维度为 $d_h = 32$的门控循环单元。上下文编码器将新息代理量 $\tilde{\boldsymbol{\nu}}_k$ 投影至隐藏空间，随后投影层将门控循环单元的输出映射到低维流形（$d_p = 16$），最后由策略头生成西格玛点权重。在训练和评估阶段，控制西格玛点扩散度的尺度参数始终固定为 $\gamma = 3$。

C. 定量结果

我们进行了蒙特卡洛分析，在训练数据集和未见过的高机动蛇形运动数据集上共开展了 N =1000个独立跟踪回合，每个回合均在长度为60个时间步的序列上进行评估。表 I 总结了位置估计的平均均方根误差（ARMSE）及其标准差。

表 I 蒙特卡洛性能基准（1000 次运行）

方法	训练方案 平均均方根误差（m）	未见机动 平均均方根误差（m）
UKF	105.0 ±129.6	196.0 ±229.9
交互多模型无迹卡尔曼滤波器	86.4 ±121.5	184.5 ±175.8
UKF*	17.8 ±14.5	49.7 ±33.9
交互多模型无迹卡尔曼滤波器*	18.5 ±17.4	58.0 ±48.8
MA-UKF（本文方法）	**6.3** ±7.3	**44.6** ±28.8

1）重尾噪声下的鲁棒性：在训练方案中，生成式运动模型与滤波器内部动力学高度吻合，基线性能主要受传感器异常的限制。标称UKF的平均均方根误差为105。0米，而优化后的UKF*达到17。8米。这种性能下降直接归因于重尾测量噪声。无迹变换的标准加权平均运算对离群值高度敏感；因此，单次闪烁测量会显著偏置后验均值，并违背隐含的高斯假设。

相比之下，MA-UKF 的平均均方根误差仅为 6.3 米；与优化后的 UKF* 相比，误差降低了 64.6%，与标称 UKF 相比则降低了 94.0%。通过分析潜在创新嵌入，网络识别出闪烁的频谱特征，即高振幅和低时间相关性。系统会做出响应

动态降低西格玛点的权重。这种机制无需脆弱的验证门控或显式阈值，即可无缝实现“软性”离群值抑制。

2) 对未见机动的泛化（分布外）：该评估方案采用一条高敏捷编织轨迹，该轨迹相对于滤波器的恒定转率预测模型存在严重的结构失配。此场景检验算法在模型信任与测量依赖之间取得平衡的能力。

如图2所示，基线滤波器无法应对未建模动态与极端噪声叠加所带来的严峻程度。无迹卡尔曼滤波器* 出现跟踪发散，完全丢失目标。交互多模型无迹卡尔曼滤波器* 在试图切换动态模型以重新捕获目标时，出现大量校正伪影以及不稳定的大幅跳变。

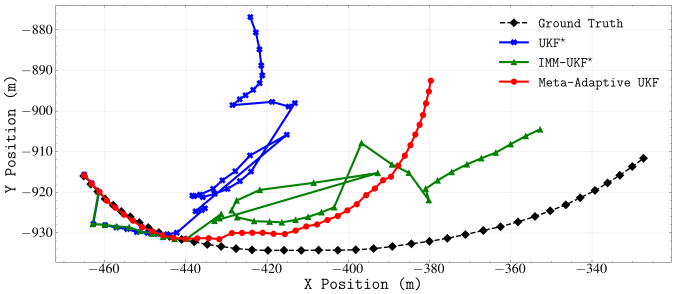

Fig. 2. 优化后的机动上的轨迹跟踪性能

优化后的无迹卡尔曼滤波器* （蓝色）出现灾难性发散，而交互多模型无迹卡尔曼滤波器* （绿色）则表现出剧烈校正。MA-UKF（红色）能在更长时间内保持稳健的跟踪连续性，并紧密贴合底层真实轨迹，尽管同时存在模型失配和严重闪烁噪声。

相反，MA-UKF 成功保留了真实轨迹的结构几何特性。它在更长时间内保持了稳定的跟踪连续性，并有效平滑了严重异常，性能分别比无迹卡尔曼滤波器* 高出 10.3%，比交互多模型无迹卡尔曼滤波器* 高出 23.1%。值得注意的是，MA-UKF 误差的标准差（± 28.8m）几乎比标称基线 UKF（± 229.9m）低 $8\times$ 。基线方法中严重的方差与图中出现的发散事件直接相关。通过调节 UT 权重，MA-UKF 构建了灵活的不确定性流形，并即时增大协方差，从而得到高度稳健的估计。

D. 学得的自适应策略分析

为阐明元学习器的决策逻辑，我们分析合成 sigma 点权重的时间演变。图 3 展示了代表性跟踪序列中，均值（$W^{(m)}$）和协方差（$W^{(c)}$）分布的完整权重轨迹集合（从 W_0 到 W_{10}）。

与标准滤波的静态参数化（其中 UT 缩放常数严格固定）相反，MA-UKF 展现出高度动态、依赖上下文的加权策略。分析揭示了两种截然不同的行为范式，体现了网络兼顾稳态精度与瞬态鲁棒性的双重能力：

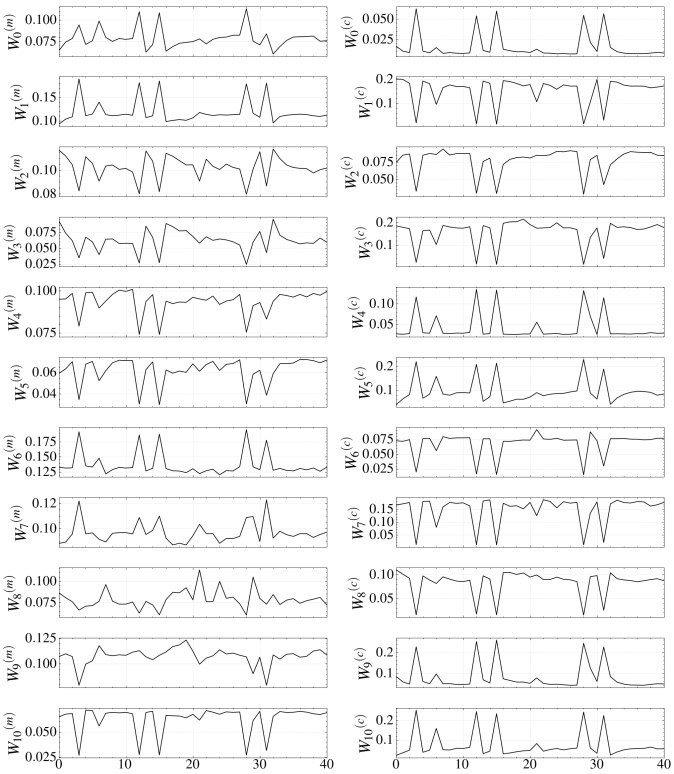

图 3. 机动期间学习到的西格玛点权重（W_0 至 W_{10}）的时间演变。整个几何形态中明显的脉冲式尖峰行为，与事件之间的连续微调制形成对比，表明策略会动态扩展和收缩局部几何扩散，以适应快速的方向变化并排除异常。

1) 连续微调制：在主要动态变化之间，整组均值与协方差权重呈现持续变化的几何结构，围绕名义基线动态波动。这表明滤波器在每个时间步都会主动且持续地补偿固定 UT 散布固有的局部线性化误差。

2) 脉冲协方差重置： 检测到未建模的动态偏移或严重的角闪烁异常后，策略会触发西格玛点权重的快速重构。如图 3 所示，该分布总体保持平坦的基线，同时触发尖锐且稀疏的尖峰，这些尖峰与机动拐点和离群事件严格相关。这种协调的脉冲式操作会注入结构不确定性，以适应未建模的加速度，从而提高卡尔曼增益的响应性，同时利用潜在上下文嵌入过滤非物理性的传感器噪声。

这些观察结果直观地证实，MA-UKF 成功地内化了一种控制策略。它并未收敛到静态最优参数，而是动态调度 11 对 UT 权重作为控制输入，持续权衡对过程与量测模型的信任程度，从而在完全连续、可微的架构中模拟出一种先进的自适应滤波器。

V. 结论与未来工作

本文提出了元自适应无迹卡尔曼滤波器（MA-UKF），这是一种将严谨的贝叶斯估计与数据驱动的元学习相结合的新型框架。通过将无迹变换的参数化构建为端到端可微优化问题，MA-UKF 学会实时动态调整其西格玛点几何。这使滤波器无需依赖人工设计的启发式方法，便能自主适应非平稳动力学和分布外工况。

通过利用循环上下文编码器将历史新息序列映射为可用于决策的潜在嵌入，所提出的架构能够有效地区分短暂的传感器异常与真实的目标机动。在重尾闪烁噪声下复杂且分布外的跟踪场景中进行的实证结果表明，与名义基线和经过严格超参数优化的无迹卡尔曼滤波器基线相比，MA-UKF分别将位置平均均方根误差降低了77.2%和10.3%，展现出稳健的泛化能力。

未来工作将重点利用真实世界的传感器数据验证MA-UKF，以评估其仿真到现实迁移性，并将可微分公式扩展至李群，以支持航空航天应用中的稳健三维位姿与方向估计。

参考文献

[1] S. J. Julier 和 J. K. Uhlmann，“卡尔曼滤波器向非线性系统的新扩展”，载于信号处理、传感器融合与目标识别 *VI* 论文集，第 3068 卷。Spie，1997 年，第 182–193 页。[2] E. A. Wan 和 R. Van Der Merwe，“用于非线性估计的无迹卡尔曼滤波器”，载于*IEEE 2000* 年自适应系统信号处理、通信与控制研讨会论文集。IEEE，2000 年，第 153–158 页。[3] R. Turner 和 C. E. Rasmussen，“无迹卡尔曼滤波中西格玛点的基于模型学习”，神经计算，第 80 卷，53，2012 年，第 47 页。–。[4] Y. Bar-Shalom、X. R. Li 和 T. Kirubarajan，用于跟踪与导航的估计：理论、算法与软件。John Wiley & Sons，2001 年。[5] A. P. Sage 和 G. W. Husa，“先验统计未知情况下的自适应滤波”，载于联合自动控制会议论文集，第 7 期，1969 年，第 760–769 页。[6] H. A. Blom 和 Y. Bar-Shalom，“具有马尔可夫切换系数系统的交互多模型算法”，*IEEE*自动控制汇刊，第 33 卷，第 8 期，第 780–783 页，1988 年。[7] N. Shlezinger、G. Revach、A. Ghosh、S. Chatterjee、S. Tang、T. Imbiriba、J. Dunik、O. Straka、P. Closas 和 Y. C. Eldar，“人工智能辅助卡尔曼滤波器：卡尔曼类算法的人工智能增强设计”，*IEEE*信号处理杂志，2025 年。[8] X. Wang 和 Y. Huang，“基于扩展卡尔曼滤波器训练循环神经网络的收敛性研究”，*IEEE*神经网络汇刊，第 22 卷，第 4 期，第 588–600 页，2011 年。[9] N. Forti、L. M. Millefiori、P. Braca 和 P. Willett，“用于机动目标跟踪的基于模型的深度学习”，载于*2023*年第*26*届国际信息融合会议（*FUSION*）。IEEE，2023 年，第 1–6 页。[10] G. Wang、P. Cheng、S. Li、Y. Chen、B. Vucetic 和 Y. Li，“用于增强室内跟踪的自监督深度状态空间模型”，载于*ICC 2025*年*IEEE*国际通信会议。IEEE，2025 年，第 692–697 页。[11] C. Jia、J. Ma 和 W. M. Kouw，“用于船舶轨迹预测并考虑不确定性的多重变分卡尔曼-GRU”，*IEEE*航空航天与电子系统汇刊，第 61 卷，第 2 期，第 3654–3667 页，2024 年。[12] D. Li、J. Zhou 和 Y. Liu，“基于循环神经网络的无迹卡尔曼滤波器，用于实时估计并补偿 MEMS 陀螺仪的随机漂移”，机械系统与信号处理，第 147 卷，文章编号 107057，2021 年。

[13] T. Haarnoja、A. Ajay、S. Levine 和 P. Abbeel，“Backprop KF：学习具有判别性的确定性状态估计器”，神经信息处理系统进展（NeurIPS），第 29 卷，2016 年。[14] R. Jonschkowski、D. Rastogi 和 O. Brock，“可微粒子滤波器：利用算法先验进行端到端学习”，载于机器人学：科学与系统论文集，美国宾夕法尼亚州匹兹堡，2018 年 6 月。[15] H. Wen、X. Chen、G. Papagiannis、C. Hu 和 Y. Li，“用于可微粒子滤波器的端到端半监督学习”，载于2021年IEEE国际机器人与自动化会议（ICRA）。IEEE，2021 年，第 5825–5831 页。[16] J. Trejbal、J. Matouˇsek 和 J. Dun´ık，“用于点质量滤波器的神经网络增强自适应网格设计”，载于2025年第26届国际喀尔巴阡控制会议（ICCC）。IEEE，2025 年，第 1–6 页。[17] G. Revach、N. Shlezinger、X. Ni、A. L. Escoriza、R. J. Van Sloun 和 Y. C. Eldar，“KalmanNet：面向部分已知动力学的神经网络辅助卡尔曼滤波”，IEEE信号处理汇刊，第 70 卷，第 1532–1547 页，2022 年。[18] J. Y. Loo、Z. Y. Ding、V. M. Baskaran、S. G. Nurzaman 和 C. P. Tan，“用于动态系统的西格玛点卡尔曼滤波器：通过优化和数据驱动方法进行非线性未知输入估计”，IEEE系统、人类与控制论汇刊：系统，第 54 卷，第 10 期，第 6068–6081 页，2024 年。[19] Z. Fan、D. Shen、Y. Bao、K. Pham、E. Blasch 和 G. Chen，“RNN-UKF：通过循环神经网络增强无迹卡尔曼滤波器的超参数自动调优”，载于2024年第27届国际信息融合会议（FUSION）。IEEE，2024 年，第 1–8 页。[20] H. Liu、Y. Chen、X. Sun、Y. Zhu、X. Wang 和 H. Gui，“通过神经网络微调高斯滤波器中的采样点”，IEEE信号处理快报，第 33 卷，第 371–375 页，2025 年。[21] Y. Duan、J. Schulman、X. Chen、P. L. Bartlett、I. Sutskever 和 P. Abbeel，“RL^2：通过慢速强化学习实现快速强化学习”，arXiv 预印本 arXiv:1611.02779，2016 年。[22] C. Finn、P. Abbeel 和 S. Levine，“用于深度网络快速适应的模型无关元学习”，载于国际机器学习会议（ICML）论文集。PMLR，2017 年，第 1126–1135 页。[23] J. X. Wang、Z. Kurth-Nelson、D. Tirumala、H. Soyer、J. Z. Leibo、R. Munos、C. Blundell、D. Kumaran 和 M. Botvinick，“学习强化学习”，arXiv 预印本 arXiv:1611.05763，2016 年。[24] N. Mishra、M. Rohaninejad、X. Chen 和 P. Abbeel，“一种简单的神经注意力元学习器”，arXiv 预印本 arXiv:1707.03141，2017 年。[25] K. Cho、B. van Merri¨enboer、C. Gulcehre、D. Bahdanau、F. Bougares、H. Schwenk 和 Y. Bengio，“使用RNN编码器–解码器学习用于统计机器翻译的短语表示”，载于2014年自然语言处理实证方法会议（EMNLP）论文集，A. Moschitti、B. Pang 和 W. Daelemans 编。卡塔尔多哈：计算语言学协会，2014 年 10 月，第 1724–1734 页。[26] L. Franceschi、M. Donini、P. Frasconi 和 M. Pontil，“基于梯度的超参数正向与反向优化”，载于国际机器学习会议（ICML）论文集。PMLR，2017 年，第 1165–1173 页。[27] A. Kloss、M. Bauza、J. Wu、J. B. Tenenbaum、A. Rodriguez 和 J. Bohg，“通过接触推理实现准确的基于视觉的操作”，载于2020年IEEE国际机器人与自动化会议（ICRA）。IEEE，2020 年，第 6738–6744 页。[28] I. Murray，“乔列斯基分解的微分”，arXiv 预印本 arXiv:1602.07527，2016 年。[29] P. Werbos，“时间反向传播：其作用及实现方法”，IEEE论文集，第 78 卷，第 10 期，第 1550–1560 页，1990 年。[30] T. Hospedales 、 A. Antoniou 、 P. Micaelli 和 A. Storkey，“神经网络中的元学习：综述”，IEEE模式分析与机器智能汇刊，第 44 卷，第 9 期，第 5149–5169 页，2021 年。[31] Z. Chen、C. Heckman、S. Julier 和 N. Ahmed，“NEES较弱？利用贝叶斯优化自动调优卡尔曼滤波器”，载于2018年第21届国际信息融合会议（FUSION），2018 年，第 1072–1079 页。[32] J. L. Ba、J. R. Kiros 和 G. E. Hinton，“层归一化”，2016 年。

[ ] 1607 06450 在线。可访问：[33] https://arxiv.org/abs/。L. Cork 和 R. Walker，“使用非线性动态模型和交互多模型无迹卡尔曼滤波器算法进行无人机传感器故障检测”，载于2007年信息、决策与控制会议。IEEE，2007 年，第 230–235 页。