

DOI:10.11918/202301052

# 一种高斯-重尾切换分布鲁棒卡尔曼滤波器

黄伟<sup>1</sup>, 付红坡<sup>1,2</sup>, 李煜<sup>1</sup>, 章卫国<sup>1</sup>

(1. 陕西省飞行控制与仿真技术重点实验室(西北工业大学), 西安 710072;

2. 信息融合技术教育部重点实验室(西北工业大学), 西安 710072)

**摘要:** 为降低实际应用中由强未知干扰和仪器故障对观测造成的影响,减轻随机和未建模干扰对系统的侵蚀,从而提升系统在非高斯噪声环境下的状态估计精度,提高滤波器的鲁棒性能,提出了一种基于高斯-重尾切换分布的鲁棒卡尔曼滤波器(Gaussian-heavy-tailed switching distribution based robust Kalman filter, GHTSRKF)。首先,通过自适应学习高斯分布和一种重尾分布之间的切换概率将噪声建模为 GHTS(Gaussian-heavy-tailed switching)分布,所设计的 GHTS 分布可以通过在线调整高斯分布和新的重尾分布之间的切换概率来对非平稳重尾噪声进行建模,具有虚拟协方差的高斯分布用于处理协方差矩阵不准确的高斯噪声。其次,引入两个分别服从 Categorical 分布与伯努利分布的辅助参数将 GHTS 分布表示为一个分层高斯形式,进一步利用变分贝叶斯方法推导了 GHTSRKF。最后,利用一个仿真场景对几种不同的 RKF(robust Kalman filters)进行了对比验证。结果表明,所提出的 GHTSRKF 算法的估计精度对初始状态的选取不敏感,精度优于其他 RKF,它的 RMSEs 最接近噪声信息准确的 KFTNC(KF with true noise covariances)的 RMSEs(root mean square errors),且当系统与量测噪声是未知时变高斯噪声时,相比于现有的滤波器, GHTSRKF 具有更好的估计性能,从而验证了 GHTSRKF 的有效性。

**关键词:** 状态估计; 非平稳重尾噪声; 自适应学习; 鲁棒滤波器; 变分贝叶斯方法

中图分类号: V249

文献标志码: A

文章编号: 0367-6234(2024)04-0012-12

## A Gaussian-heavy-tailed switching distribution robust Kalman filter

HUANG Wei<sup>1</sup>, FU Hongpo<sup>1,2</sup>, LI Yu<sup>1</sup>, ZHANG Weiguo<sup>1</sup>

(1. Shaanxi Provincial Key Laboratory of Flight Control and Simulation Technology (Northwestern Polytechnical University), Xi'an 710072, China; 2. Key Laboratory of Information Fusion Technology (Northwestern Polytechnical University), Ministry of Education, Xi'an 710072, China)

**Abstract:** To mitigate the influence of strong unknown disturbances and instrument faults on observations in practical applications, and to alleviate the degradation caused by random and unmodeled interferences on the system, so as to improve the state estimation accuracy of the system in non-Gaussian noise environment and the robustness of the filter, a Gaussian-heavy-tailed switching distribution based robust Kalman filter (GHTSRKF) is proposed. Firstly, the noises are modeled as a GHTS (Gaussian-heavy-tailed switching) distribution by adaptively learning the switching probability between the Gaussian distribution and the newly designed heavy-tailed distribution. The designed GHTS distribution can model non-stationary heavy tail noise by adjusting the switching probability between the Gaussian distribution and the new heavy-tailed distribution online. The Gaussian distribution with a virtual covariance is used to deal with Gaussian noise with inaccurate covariance matrix. Secondly, two auxiliary parameters following the category distribution and the Bernoulli distribution are introduced to express the GHTS distribution as a hierarchical Gaussian form. Furthermore, the GHTSRKF is derived by utilizing the variational Bayesian method. Finally, a simulation scenario is used to compare and verify several different robust Kalman filters (RKF). The results show that the accuracy of the proposed GHTSRKF algorithm is insensitive to the selection of initial state and exhibits higher estimation accuracy compared to other RKF. Its root mean square errors (RMSEs) are closest to those of KF with true noise covariances (KFTNC) with accurate noise information. Compared with existing filters, GHTSRKF has better estimation performance when the system and measurement noise are unknown time-varying Gaussian noise, thus verifying the effectiveness of GHTSRKF.

**Keywords:** state estimation; non-stationary heavy-tailed noises; adaptive learning; robust filter; variational Bayesian method

收稿日期: 2023-01-15; 录用日期: 2023-03-08; 网络首发日期: 2024-03-09

网络首发地址: <https://link.cnki.net/urlid/23.1235.t.20240307.1120.002>

基金项目: 国家自然科学基金(62073266)

作者简介: 黄伟(1994—),男,博士研究生;章卫国(1956—),男,教授,博士生导师

通信作者: 付红坡, 17610828038@163.com

在组合导航、目标跟踪、飞行器姿态控制、故障诊断等众多工程领域中,系统的状态估计是关键问题。在状态估计方法中,卡尔曼滤波器(Kalman filter, KF)因其简洁性、高效性与最优性被广泛应用于上述领域<sup>[1-6]</sup>。但是,在实际应用中,观测常常受到强未知干扰和仪器故障的影响,系统本身也会受到随机和未建模干扰的侵蚀,这些都容易造成未知异常值的出现。这种情况下,过程与量测噪声往往是非高斯重尾分布,这违背了KF的噪声是精确高斯的假设,从而破坏了滤波器的准确性和可靠性<sup>[7-10]</sup>。对于这一问题,许多优秀的研究者已经提出了很多鲁棒卡尔曼滤波器(robust Kalman filters, RKF)。

虽然 Huber-RKFs<sup>[11-12]</sup>和最大/最小相关熵准则 RKFs<sup>[13-15]</sup>可以在一定程度上降低重尾噪声的负面影响,但由于没有充分利用该噪声固有的重尾特性与不能适用于复杂的噪声环境限制了它们的性能。针对这一问题,学生 t KF(Student's t KF, STKF)<sup>[16]</sup>和鲁棒学生 t KF(robust Student's t KF, RSTKF)<sup>[17]</sup>先后被提出,其中状态后验的概率密度函数(probability density function, PDF)被建模为学生 t 分布,并通过变分贝叶斯(variational Bayesian, VB)推理估计状态和噪声统计量。为进一步提高滤波器的鲁棒性能,基于野值鲁棒学生 t KF(outlier-robust Student's t KF, ORSTKF)<sup>[18]</sup>和基于重尾混合分布的 RKF(heavy-tailed mixture RKF, HTMRKF)<sup>[19]</sup>被设计。虽然上述提到的 RKFs 在平稳重尾噪声环境中表现良好,但是如果存在由动态异常值引起的非平稳重尾过程和测量噪声,即噪声有时是高斯噪声,有时是不同类型的重尾分布,上述 RKFs 的性能会下降。例如在自主水下航行器协同定位中,由于传感器的离群值、声通道的多径效应以及水下环境的变化,实际状态和测量噪声可能是非平稳非高斯分布的<sup>[20]</sup>。还有在无人潜航器的惯性导航/多普勒组合导航实际实验中,也出现了类似的噪声环境<sup>[21]</sup>。因此在设计鲁棒滤波器时需考虑这样的噪声环境。

考虑到上述问题,一种新的高斯-学生 t 混合分布(Gaussian-student's t mixture distribution, GSTM)<sup>[22]</sup>被构造以建模非平稳重尾噪声,该算法通过自适应学习高斯分布和学生 t 分布之间的混合概率来适应噪声的非平稳性。并通过引入伯努利随机变量,将所提出的 GSTM 分布表示为分层高斯形式,并在此基础上构造了新的分层线性高斯状态空间模型以推导一种新的基于 GSTM 分布的 RKF(GSTM RKF),其中状态向量、辅助随机变量、伯努利随机变量和混合概率用 VB 方法联合推导。仿真结果表明,在非平稳重尾噪声情况下,该滤波器比现有的标准 KF 和 ORSTKF 具有更好的估计精度。这是由于混合参数

克服了自由度参数自适应的困难,能更好地模拟非平稳重尾分布噪声。然而该滤波器的噪声尺度矩阵必须是预先已知的。在文献[22]工作的基础上,基于两种高斯分布混合(mixture of two Gaussian distribution)的 RKF(M2GRKF)<sup>[23]</sup>被提出。但实际上,由于不同类型的异常值随机出现,非平稳重尾噪声的特征是难以描述的,这使得上述 RKFs 很难预先选择尺度矩阵和利用单个伽马分布的粗略先验信息来近似重尾特征。

本文以应对上述难题为目标,提出了一种基于高斯-重尾切换分布的鲁棒卡尔曼滤波器(Gaussian-heavy-tailed switching distribution based robust Kalman filter, GHTSRKF)。所设计的 GHTS 分布可以通过在线调整高斯分布和新的重尾分布之间的切换概率来对非平稳重尾噪声进行建模。具有虚拟协方差的高斯分布用于处理协方差矩阵不准确的高斯噪声。重尾分布由高斯分布和混合伽马分布构成,前者利用一个尺度参数来描述重尾噪声;后者可以根据精确先验对尺度参数进行建模,所使用的精确先验是通过自适应学习混合概率向量从一组粗略的先验信息中提取的。本文首先建立一种新的 GHTS 分布对实际情况可能出现的非平稳重尾噪声进行建模。其次引入两个分别满足分类分布和伯努利分布的随机变量将 GHTS 分布 PDF 转换为分层高斯形式以便于算法的进一步推导。最后利用 VB 方法推导出 GHTSRKF,该滤波器中一步预测和测量似然 PDFs 被建模为 GHTS 分布,并采用固定点迭代联合估计状态和模型参数。

## 1 问题描述

离散时间状态空间模型如下:

$$\begin{cases} \mathbf{x}_t = \mathbf{F}_{t-1} \mathbf{x}_{t-1} + \boldsymbol{\omega}_{t-1} \\ \mathbf{z}_t = \mathbf{H}_t \mathbf{x}_t + \mathbf{v}_t \end{cases} \quad (1)$$

式中:  $\mathbf{x}_t \in \mathbf{R}^{n_x}$ ,  $\mathbf{z}_t \in \mathbf{R}^{m_z}$  分别为状态向量与量测向量,  $\mathbf{F}_{t-1}$ 、 $\mathbf{H}_t$  分别为状态转移矩阵与量测矩阵,  $\boldsymbol{\omega}_{t-1}$ 、 $\mathbf{v}_t$  分别为系统与量测噪声向量,且协方差矩阵分别为  $\mathbf{Q}_{t-1}$  与  $\mathbf{R}_t$ ,同时,假设状态与两种噪声相互独立。

KF 可以根据给定的状态空间模型与量测  $\mathbf{z}_{1:t}$  估计出系统状态  $\mathbf{x}_t$ 。在卡尔曼理论中,当状态空间模型是线性的且系统与量测噪声是准确已知的高斯分布时, KF 的估计是最优的。然而实际情况中,系统本身与观测传感器由于变化的环境和自身的原因,过程与量测中会出现未知的动态野值,导致系统的过程和测量噪声是非平稳重尾分布,这违背了卡尔曼理论中精确高斯分布的假设,严重影响到了滤波的估计精度和可靠性。

考虑到上述现实问题,本文以降低 KF 对精确先验噪声统计的依赖,并改善在复杂噪声,如重尾噪声、非平稳重尾噪声等环境下的状态估计性能为目标,研究一种应用更广、性能更好的鲁棒滤波器。

## 2 高斯-重尾切换分布(GHTS)

本文首先引入了一种新的重尾分布,也称为高斯-混合伽马(Gaussian-mixing Gamma, GMG)分布以建模重尾噪声<sup>[10]</sup>。基于混合概率向量  $\xi$  的服从 GMG 分布的随机向量  $y$  的 PDF 表示为

$$p(y|\xi) = N(y; \lambda, \mathbf{A}/\kappa) \sum_{m=1}^M \xi_m G(\kappa; e_m, f_m) \quad (2)$$

式中:  $N(\cdot; \lambda, \mathbf{A})$  表示均值为  $\lambda$  和协方差矩阵为  $\mathbf{A}$  的高斯分布,  $G(\cdot; e, f)$  表示形状参数为  $e$  和速率参数为  $f$  的伽马分布,  $\kappa$  为尺度参数,  $\xi_m$  为第  $m$  个伽马分布的混合系数,  $M$  为伽马分布的数量,  $\xi = [\xi_1, \xi_2, \dots, \xi_M]$  为混合概率向量且  $\sum_{m=1}^M \xi_m = 1$  被用于产生不同的伽马分布的混合来建模不同情况的重尾噪声。

文献[24]考虑了贝叶斯框架下高斯混合分布的标准共轭先验,利用狄利克雷分布对混合系数进行建模。基于该方法,本文假定混合概率向量  $\xi$  服从如下狄雷克利分布:

$$p(\xi) = \text{Dir}(\xi; \mathbf{w}, m) \quad (3)$$

式中,向量  $\mathbf{w} = [w_1, w_2, \dots, w_m]$  且  $w_m > 0$  是狄利克雷分布的集中参数。

本文假设变量  $\xi_m$  的期望为  $\bar{w}_m = E[\xi_m]$ , 并利用式(2)、(3)可以将 GMG 分布表示为

$$p(y) = \int p(y|\xi) p(\xi) d\xi = N(y; \lambda, \mathbf{A}/\kappa) \sum_{m=1}^M \bar{w}_m G(\kappa; e_m, f_m) \quad (4)$$

为准确描述非平稳重尾噪声,设计了一种新的 GHTS 分布。服从该分布的随机向量  $y$  的 PDF 为

$$p(y|\mu) = \mu N(y; \lambda, \mathbf{A}) + (1 - \mu) \text{GMG}(y, \lambda, \mathbf{A}, e, f) \quad (5)$$

文献[10]设计了贝叶斯框架下实现了两种不同噪声分布之间的切换,并利用贝塔分布建模两种分布之间的切换概率。基于该结果,本文的切换概率  $\mu$  也服从以下贝塔分布:

$$p(\mu) = B(\mu; v, 1 - v) \quad (6)$$

式中:  $B(\cdot; v, 1 - v)$  为形状参数为  $v \in [0, 1]$  的贝塔分布,  $v$  为先验切换概率,则  $v = E[\mu]$ 。

根据式(5)、(6), GHTS 的 PDF 表示为

$$p(y) = \int p(y|\mu) p(\mu) d\mu = v N(y; \lambda, \mathbf{A}) + (1 - v) \text{GMG}(y, \lambda, \mathbf{A}, e, f) \quad (7)$$

引入一个伯努利参数  $\zeta$  和一个服从分类分布的向量  $\rho = [\rho_1, \rho_2, \dots, \rho_M]$ , 将 GHTS 分布的 PDF 表示为

$$p(y) = \sum_{\zeta=0}^1 \iiint_0^{+\infty} [N(y; \lambda, \mathbf{A})]^\zeta \times [N(y; \lambda, \mathbf{A}/\kappa)]^{(1-\zeta)} \times \prod_{m=1}^M [G(\kappa; e_m, f_m)]^{\rho_m} \times p(\zeta|\mu) p(\mu) p(\rho|\xi) p(\xi) d\zeta d\kappa d\mu d\xi, \zeta \in \{0, 1\} \quad (8)$$

其中  $\zeta$ 、 $\rho$  的 PDFs 为

$$\begin{cases} p(\zeta|\mu) = \text{Bern}(\zeta; \mu) = \mu^\zeta (1 - \mu)^{1-\zeta} \\ p(\rho|\xi) = \text{Cat}(\rho; \xi, M) \end{cases}$$

式中:  $\text{Bern}(\cdot)$  为伯努利分布,  $\text{Cat}(\cdot)$  为 Categorical 分布。式(8)中的混合概率向量  $\xi$  和切换概率  $\mu$  已经在式(3)和式(5)中被假设为狄利克雷和贝塔分布。

由于先验噪声统计不准确与野值干扰使得尺度矩阵  $\mathbf{A}$  不太准确,假设其服从<sup>[23]</sup>:

$$p(\mathbf{A}) = \text{IW}(\mathbf{A}; \mathbf{V}, \sigma) \quad (9)$$

式中  $\text{IW}(\cdot; \mathbf{V}, \sigma)$  为逆尺度矩阵  $\mathbf{V}$  和自由度参数  $\sigma$  的逆威沙特分布。

利用式(3)~式(9)可将 GHTS 分布表示为

$$\begin{cases} p(y|\zeta, \kappa, \mathbf{A}) = [N(y; \lambda, \mathbf{A})]^\zeta [N(y; \lambda, \mathbf{A}/\kappa)]^{(1-\zeta)} \\ p(\kappa|\rho) = \prod_{m=1}^M [G(\kappa; e_m, f_m)]^{\rho_m} \\ p(\zeta|\mu) = \text{Bern}(\zeta; \mu) \\ p(\rho|\xi) = \text{Cat}(\rho; \xi, M) \\ p(\xi) = \text{Dir}(\xi; \mathbf{w}, M) \\ p(\mu) = B(\mu; v, 1 - v) \\ p(\mathbf{A}) = \text{IW}(\mathbf{A}; \mathbf{V}, \sigma) \end{cases} \quad (10)$$

接下来,基于式(10)和 VB 方法将推导出一种新的 GHTSRKF。

**注 1** 针对实际工程领域中可能因传感器故障、环境干扰等造成的野值引发的重尾噪声情形,文献[17-18]中的高斯伽马分布/学生 t 被广泛用于建模该类型噪声。进一步,为了更精确地建模该类型噪声,文献[10]中设计了一种高斯-混合伽马分布,即式(2)中的 GMG 分布来应对复杂的重尾噪声。它是基于多个不同参数的伽马分布的一组粗糙先验得到的精确先验来描述噪声尺度参数。因此本文利用 GMG 分布来描述重尾噪声。但考虑到 GMG 分布在无异常值情况下的性能受限,本文开发了如式(7)所示的 GHTS 分布,这可以扩展 GMG 的应用范围,提高其适用性与估计性能。

**注 2** 在文献[23]中,设计了一种基于两个高斯分布混合的鲁棒滤波器,本文的 GHTS 分布与其相似。但相比于文献[23]中的 M2G 分布,在系统

噪声与量测噪声是未知重尾分布时, GHTS 分布中的混合伽马分布的设计能够预先提供一组粗略先验信息, 并通过混合概率向量的自适应学习提取出一个准确的先验, 这使得提出的滤波器能够对不同的重尾噪声进行更加准确描述, 鲁棒性也更强。而 M2G 分布中只能有一个协方差较大的高斯分布以应对重尾噪声, 因此在应对重尾分布噪声时其性能略差于 GHTS 分布。这将在仿真中进一步验证。

**注3** 在式(2)中, 由于混合伽马分布中元素的数量直接决定了 GHTS 分布的性能, 其选择应慎重。当  $M$  较小时, 由于缺乏足够的先验噪声信息, 滤波器的估计精度较差。随着  $M$  的增加, 滤波器的估计精度会提高, 而学习的参数越多, 算法的计算量就越

$$\begin{cases} p(\mathbf{x}_t | \mathbf{z}_{1:t-1}) = \sum_{y_t=0}^1 \iiint \int_0^{+\infty} [N(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t-1}, \boldsymbol{\Sigma}_t)]^{y_t} \times [N(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t-1}, \boldsymbol{\Sigma}_t / \sigma_t)]^{1-y_t} \prod_{i=1}^I [G(\sigma_t; a_{0,i}, b_{0,i})]^{\varepsilon_{t,i}} \times \\ \quad \text{Bern}(y_t; \delta_t) B(\delta_t; k_0, 1 - k_0) \text{Cat}(\boldsymbol{\varepsilon}_t; \boldsymbol{\theta}_t, I) \times \text{Dir}(\boldsymbol{\theta}_t; \hat{\mathbf{e}}_{t|t-1}, I) \text{IW}(\boldsymbol{\Sigma}_t; s_{t|t-1}, \mathbf{S}_{t|t-1}) d\mathbf{y}_t d\sigma_t d\delta_t d\boldsymbol{\theta}_t d\boldsymbol{\Sigma}_t \\ p(\mathbf{z}_t | \mathbf{x}_t) = \sum_{r_t=0}^1 \iiint \int_0^{+\infty} [N(\mathbf{z}_t; \mathbf{H}_t \mathbf{x}_t, \mathbf{R}_t)]^{r_t} \times [N(\mathbf{z}_t; \mathbf{H}_t \mathbf{x}_t, \mathbf{R}_t / \pi_t)]^{1-r_t} \prod_{l=1}^L [G(\pi_t; c_{0,l}, d_{0,l})]^{\eta_{t,l}} \times \\ \quad \text{Bern}(r_t; \lambda_t) B(\lambda_t; h_0, 1 - h_0) \times \text{Cat}(\boldsymbol{\eta}_t; \boldsymbol{\tau}_t, L) \times \text{Dir}(\boldsymbol{\tau}_t; \hat{\mathbf{g}}_{t|t-1}, L) \text{IW}(\mathbf{R}_t; u_{t|t-1}, \mathbf{U}_{t|t-1}) d\mathbf{r}_t d\pi_t d\lambda_t d\boldsymbol{\tau}_t d\mathbf{R}_t \end{cases} \quad (11)$$

式中:  $y_t, r_t$  为伯努利变量,  $\sigma_t, \pi_t$  为尺度参数,  $\boldsymbol{\varepsilon}_t, \boldsymbol{\eta}_t$  为分类分布向量,  $\delta_t, \lambda_t$  为切换概率,  $\boldsymbol{\theta}_t, \boldsymbol{\tau}_t$  为混合概率向量,  $\boldsymbol{\Sigma}_t, \mathbf{R}_t$  为尺度矩阵,  $I, L$  为伽马分布的数量。

均值向量  $\hat{\mathbf{x}}_{t|t-1}$  为

$$\hat{\mathbf{x}}_{t|t-1} = \mathbf{F}_t \hat{\mathbf{x}}_{t-1|t-1} \quad (12)$$

先验集中参数向量  $\mathbf{e}_{t|t-1}, \mathbf{g}_{t|t-1}$ , 先验自由度参数  $s_{t|t-1}, u_{t|t-1}$ , 先验逆尺度矩阵  $\mathbf{S}_{t|t-1}, \mathbf{U}_{t|t-1}$  分别为

$$\begin{cases} \mathbf{e}_{t|t-1} = \rho \hat{\mathbf{e}}_{t-1|t-1} \\ \mathbf{g}_{t|t-1} = \rho \hat{\mathbf{g}}_{t-1|t-1} \\ s_{t|t-1} = m_t \\ \mathbf{S}_{t|t-1} = m_t \tilde{\mathbf{P}}_{t|t-1} \\ u_{t|t-1} = \rho \hat{u}_{t-1|t-1} \\ \mathbf{U}_{t|t-1} = \rho \hat{\mathbf{U}}_{t-1|t-1} \end{cases} \quad (13)$$

式中:  $\rho \in (0, 1]$  为遗忘因子,  $m_t$  为调节参数, 这两个参数的作用将在后面分析,  $\tilde{\mathbf{P}}_{t|t-1}$  为虚拟预测误差协方差, 由下式可得

$$\tilde{\mathbf{P}}_{t|t-1} = \mathbf{F}_t \mathbf{P}_{t-1|t-1} \mathbf{F}_t^T + \tilde{\mathbf{Q}}_{t-1} \quad (14)$$

式中  $\tilde{\mathbf{Q}}_{t-1}$  为预选的虚拟系统噪声协方差。

根据式(10)、(11), 一步预测与量测似然 PDFs 的分层高斯形式可以分别表示为

大。但是, 如果  $M$  过大, 对估计结果的影响将非常有限, 甚至一个或多个混合元会陷入过拟合, 这将大大降低估计的精度。而且, 较大的  $M$  也会导致较大的计算复杂度, 不利于算法的实时性估计。综上所述, 在实际应用中, 为了平衡估计精度和计算效率, 混合数量  $M$  的设置应灵活适度。

### 3 基于 GHTS 分布的鲁棒卡尔曼滤波器 (GHTSRKF)

#### 3.1 分层高斯状态空间模型

对于式(1)所示的状态空间模型, 当过程与量测噪声是非平稳重尾噪声时, 本文将一步预测与量测似然 PDFs 建模为 GHTS 分布:

$$\begin{cases} p(\mathbf{x}_t | \mathbf{z}_{1:t-1}, y_t, \boldsymbol{\varepsilon}_t) = [N(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t-1}, \boldsymbol{\Sigma}_t)]^{y_t} [N(\mathbf{x}_t; \hat{\mathbf{x}}_{t|t-1}, \boldsymbol{\Sigma}_t / \sigma_t)]^{1-y_t} \times \\ \quad \prod_{i=1}^I [G(\sigma_t; a_{0,i}, b_{0,i})]^{\varepsilon_{t,i}} \\ p(\boldsymbol{\Sigma}_t) = \text{IW}(\boldsymbol{\Sigma}_t; s_{t|t-1}, \mathbf{S}_{t|t-1}) \\ p(y_t | \delta_t) = \text{Bern}(y_t; \delta_t), p(\delta_t) = B(\delta_t; k_0, 1 - k_0) \\ p(\boldsymbol{\varepsilon}_t | \boldsymbol{\theta}_t) = \text{Cat}(\boldsymbol{\varepsilon}_t; \boldsymbol{\theta}_t, I), p(\boldsymbol{\theta}_t) = \text{Dir}(\boldsymbol{\theta}_t; \hat{\mathbf{e}}_{t|t-1}, I) \\ p(\mathbf{z}_t | \mathbf{x}_t, r_t, \boldsymbol{\eta}_t) = [N(\mathbf{z}_t; \mathbf{H}_t \mathbf{x}_t, \mathbf{R}_t)]^{r_t} [N(\mathbf{z}_t; \mathbf{H}_t \mathbf{x}_t, \mathbf{R}_t / \pi_t)]^{1-r_t} \times \\ \quad \prod_{l=1}^L [G(\pi_t; c_{0,l}, d_{0,l})]^{\eta_{t,l}} \\ p(\mathbf{R}_t) = \text{IW}(\mathbf{R}_t; u_{t|t-1}, \mathbf{U}_{t|t-1}) \\ p(r_t | \lambda_t) = \text{Bern}(r_t; \lambda_t), p(\lambda_t) = B(\lambda_t; h_0, 1 - h_0) \\ p(\boldsymbol{\eta}_t | \boldsymbol{\tau}_t) = \text{Cat}(\boldsymbol{\eta}_t; \boldsymbol{\tau}_t, L), p(\boldsymbol{\tau}_t) = \text{Dir}(\boldsymbol{\tau}_t; \hat{\mathbf{g}}_{t|t-1}, L) \end{cases} \quad (15)$$

式(12)~式(15)构成了一个分层高斯状态空间模型, 其模型如图1所示。然后, 进一步利用 VB 方法推导出 GHTSRKF。

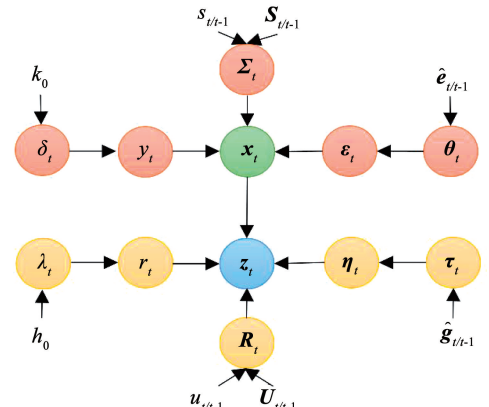


图1 基于 GHTS 分布的分层高斯状态空间模型

Fig. 1 Graphical model for hierarchical Gaussian state-space model based on GHTS distribution

### 3.2 后验 PDFs 的联合 VB 推理

正如式(15)所示,由于模型参数与状态向量之间存在相互耦合关系,通过解析的方式获取状态向量的后验估计是很麻烦的。因此为了估计变量集合  $\Theta$ , 本文利用 VB 方法通过最小化 Kullback-Leibler 发散代价函数实现对真实联合后验 PDF  $p(\Theta | z_{1:k})$  的最优近似<sup>[25]</sup>。使用的近似后验 PDF  $q(\Delta)$  满足:

$$\log q(\Delta) = E_{\Theta^{(-\Delta)}} [\log p(\Theta, z_{1:t})] + c_{\Delta} \quad (16)$$

式中:  $E[\cdot]$  为期望计算, 变量集合  $\Theta = \{x_t, \sigma_t, \pi_t, y_t, r_t, \delta_t, \lambda_t, \varepsilon_t, \eta_t, \theta_t, \tau_t, \Sigma_t, R_t\}$ ,  $\Theta^{(-\Delta)}$  为  $\Theta$  的一个子集, 其满足  $\{\Delta\} \cup \Theta^{(-\Delta)} = \Theta$ ,  $c_{\Delta}$  为独立于  $\Delta$  的常数。

考虑到式(15)中变分参数之间的耦合, 本文采用固定点迭代法对式(16)进行迭代求解以计算  $q(\Delta)$ 。在  $(j+1)$  次更新时, 利用  $q^{(j)}(\Theta^{(-\Delta)})$  将  $q(\Delta)$  更新为  $q^{(j+1)}(\Delta)$ 。根据贝叶斯理论, 真实联合 PDF 分解为

$$p(\Theta, z_{1:t}) = p(z_{1:t-1})p(x_t | z_{1:t-1}, y_t, \varepsilon_t)p(z_t | x_t, r_t, \eta_t) \times p(y_t | \delta_t)p(r_t | \lambda_t)p(\delta_t)p(\lambda_t)p(\varepsilon_t | \theta_t)p(\eta_t | \tau_t)p(\theta_t) \times p(\tau_t)p(\Sigma_t)p(R_t) \quad (17)$$

对式(17)进行对数运算可得

$$\begin{aligned} \log p(\Theta, z_{1:t}) = & -0.5y_t[\log |\Sigma_t| - (x_t - \hat{x}_{t|t-1})^T \Sigma_t^{-1} (x_t - \hat{x}_{t|t-1})] + \\ & 0.5(1-y_t)[n_x \log \sigma_t - \log |\Sigma_t| - \sigma_t(x_t - \hat{x}_{t|t-1})^T \Sigma_t^{-1} (x_t - \hat{x}_{t|t-1})] + \\ & \sum_{i=1}^I [(a_{0,i} - 1) \log \sigma_t - b_{0,i} \sigma_t] \varepsilon_{t,i} + \sum_{i=1}^I (\varepsilon_{t,i} + \hat{\varepsilon}_{t-1,i} - 1) \log \theta_{t,i} + \\ & y_t \log \delta_t + (1-y_t) \log(1-\delta_t) + (k_0 - 1) \log \delta_t - k_0 \log(1-\delta_t) - \\ & 0.5(s_{t|t-1} + 1) \log |\Sigma_t| - 0.5 \text{tr}(S_{t|t-1} \Sigma_t^{-1}) - 0.5r_t[\log |R_t| - \\ & (z_t - H_t x_t)^T R_t^{-1} (z_t - H_t x_t)] + 0.5(1-r_t)[n_z \log \pi_t - \log |R_t| - \\ & \pi_t(z_t - H_t x_t)^T R_t^{-1} (z_t - H_t x_t)] + \sum_{l=1}^L [(c_{0,l} - 1) \log \pi_t - \\ & d_{0,l} \pi_t] \eta_{t,l} + \sum_{l=1}^L (\eta_{t,l} + \hat{g}_{t-1,l} - 1) \log \tau_{t,l} + r_t \log \lambda_t + (1-r_t) \log(1-\lambda_t) + \\ & (h_0 - 1) \log \lambda_t - h_0 \log(1-\lambda_t) - 0.5(u_{t|t-1} + 1) \log |R_t| - 0.5 \text{tr}(U_{t|t-1} R_t^{-1}) \end{aligned} \quad (18)$$

本文利用固定点迭代法对系统状态与模型参数的近似后验  $q(\Delta)$  进行迭代求解<sup>[25]</sup>。

**命题 1** 该命题是在噪声尺度参数、切换概率等未知参数与噪声协方差矩阵  $\{\sigma_t, \pi_t, y_t, r_t, \delta_t, \lambda_t, \varepsilon_t, \eta_t, \theta_t, \tau_t, \Sigma_t, R_t\}$  保持不变的情况下实现状态  $x_t$  的后验估计, 即在  $\Delta = x_t$  的同时将式(17)代入式(16), 然后近似后验 PDF  $q^{(j+1)}(x_t)$  可以被迭代更新为以下高斯分布:

$$q^{(j+1)}(x_t) = N(x_t; \hat{x}_{t|t}^{(j+1)}, P_{t|t}^{(j+1)}) \quad (19)$$

式中, 后验状态向量  $\hat{x}_{t|t}^{(j+1)}$  和后验协方差  $P_{t|t}^{(j+1)}$  由标准 KF 的量测更新得到:

$$\begin{cases} K_t^{(j+1)} = \hat{P}_{t|t-1}^{(j+1)} H_t^T (H_t P_{t|t-1} H_t^T + \hat{R}_t^{(j+1)})^{-1} \\ \hat{x}_{t|t}^{(j+1)} = \hat{x}_{t|t-1} + K_t^{(j+1)} (z_t - H_t \hat{x}_{t|t-1}) \\ P_{t|t}^{(j+1)} = (I_n - K_t^{(j+1)} H_t) \hat{P}_{t|t-1}^{(j+1)} \end{cases} \quad (20)$$

式中, 估计的状态预测误差协方差  $\hat{P}_{t|t-1}^{(j+1)}$  和量测噪声协方差  $\hat{R}_t^{(j+1)}$  分别为

$$\begin{cases} \hat{P}_{t|t-1}^{(j+1)} = \frac{\{E^{(j)}[\Sigma_t^{-1}]\}^{-1}}{(1 - E^{(j)}[y_t])E^{(j)}[\sigma_t] + E^{(j)}[y_t]} \\ \hat{R}_t^{(j+1)} = \frac{\{E^{(j)}[R_t^{-1}]\}^{-1}}{(1 - E^{(j)}[r_t])E^{(j)}[\pi_t] + E^{(j)}[r_t]} \end{cases} \quad (21)$$

**推导 1** 将式(16)代入式(18)有

$$\begin{aligned} \log q^{(j+1)}(x_t) = & -0.5\{E^{(j)}[y_t] + \\ & (1 - E^{(j)}[y_t])E^{(j)}[\sigma_t]\} \times \\ & (x_t - \hat{x}_{t|t-1})^T \Sigma_t^{-1} (x_t - \hat{x}_{t|t-1}) - \\ & 0.5\{E^{(j)}[r_t] + (1 - E^{(j)}[r_t]) \times \\ & E^{(j)}[\pi_t]\} (z_t - H_t x_t)^T R_t^{-1} (z_t - H_t x_t) + c_{x_t} \end{aligned} \quad (22)$$

由贝叶斯理论可以得到命题 1 中式(19)~式(21)。

**命题 2** 该命题是在状态向量、伯努利参数等未知参数与噪声协方差矩阵  $\{x_t, y_t, r_t, \delta_t, \lambda_t, \varepsilon_t, \eta_t, \theta_t, \tau_t, \Sigma_t, R_t\}$  保持不变的情况下实现噪声尺度参数  $\sigma_t$  和  $\pi_t$  的变分近似估计, 分别使  $\Delta = \sigma_t$  和  $\Delta = \pi_t$  并将式(17)代入式(16), 然后近似后验 PDFs  $q^{(j+1)}(\sigma_t)$  和  $q^{(j+1)}(\pi_t)$  可以被迭代更新为以下伽马分布:

$$\begin{cases} q^{(j+1)}(\sigma_t) = G(\sigma_t; \hat{\alpha}_{1,t}^{(j+1)}, \hat{\alpha}_{2,t}^{(j+1)}) \\ q^{(j+1)}(\pi_t) = G(\pi_t; \hat{\beta}_{1,t}^{(j+1)}, \hat{\beta}_{2,t}^{(j+1)}) \end{cases} \quad (23)$$

式中估计的形状参数  $\hat{\alpha}_{1,t}^{(j+1)}$ 、 $\hat{\beta}_{1,t}^{(j+1)}$  和估计的速率参数  $\hat{\alpha}_{2,t}^{(j+1)}$ 、 $\hat{\beta}_{2,t}^{(j+1)}$  分别表示为

$$\begin{cases} \hat{\alpha}_{1,t}^{(j+1)} = 0.5n_x(1 - E^{(j)}[y_t]) + E^{(j)}[\varepsilon_t]a_0^T \\ \hat{\alpha}_{2,t}^{(j+1)} = 0.5(1 - E^{(j)}[y_t])\text{tr}(\Psi_t^{(j+1)} E^{(j)}[\Sigma_t^{-1}]) + E^{(j)}[\varepsilon_t]b_0^T \\ \hat{\beta}_{1,t}^{(j+1)} = 0.5n_z(1 - E^{(j)}[r_t]) + E^{(j)}[\eta_t]c_0^T \\ \hat{\beta}_{2,t}^{(j+1)} = 0.5(1 - E^{(j)}[r_t])\text{tr}(\Xi_t^{(j+1)} E^{(j)}[R_t^{-1}]) + E^{(j)}[\eta_t]d_0^T \end{cases} \quad (24)$$

辅助矩阵  $\Psi_t^{(j+1)}$ 、 $\Xi_t^{(j+1)}$  分别为

$$\begin{cases} \Psi_t^{(j+1)} = E^{(j+1)}[(x_t - F_t x_{t-1})(x_t - F_t x_{t-1})^T] \\ \Xi_t^{(j+1)} = E^{(j+1)}[(z_t - H_t x_t)(z_t - H_t x_t)^T] \end{cases} \quad (25)$$

**推导 2** 将式(16)代入式(18)有

$$\begin{cases} \log q^{(j+1)}(\sigma_t) = (0.5n_x(1-E^{(j)}[y_t]) + E^{(j)}[\epsilon_t]d_0^T - 1) \times \\ \quad \log \sigma_t - 0.5(1-E^{(j)}[y_t])\text{tr}(\Psi_t^{(j+1)}E^{(j)}[\Sigma_t^{-1}]) + \\ \quad E^{(j)}[\epsilon_t]b_0^T\sigma_t + c_{\sigma_t} = (\hat{\alpha}_{1,t}^{(j+1)} - 1)\log \sigma_t - \hat{\alpha}_{2,t}^{(j+1)}\sigma_t + c_{\sigma_t} \\ \log q^{(j+1)}(\pi_t) = (0.5n_z(1-E^{(j)}[r_t]) + E^{(j)}[\eta_t]c_0^T - 1) \times \\ \quad \log \pi_t - 0.5(1-E^{(j)}[r_t])\text{tr}(\Xi_t^{(j+1)}E^{(j)}[R_t^{-1}]) + \\ \quad E^{(j)}[\eta_t]d_0^T\pi_t + c_{\pi_t} = (\hat{\beta}_{1,t}^{(j+1)} - 1)\log \pi_t - \hat{\beta}_{2,t}^{(j+1)}\pi_t + c_{\pi_t} \end{cases} \quad (26)$$

对式(26)两边进行对数操作,并由伽马分布的定义可以得到命题2中的式(23)~式(25)。

**命题3** 该命题是在状态向量、噪声尺度参数等未知参数与噪声协方差矩阵 $\{\mathbf{x}_t, \sigma_t, \pi_t, \delta_t, \lambda_t, \epsilon_t, \eta_t, \theta_t, \tau_t, \Sigma_t, R_t\}$ 保持不变的情况下实现伯努利参数 $y_t$ 和 $r_t$ 的变分近似估计,分别使 $\Delta = y_t$ 和 $\Delta = r_t$ 并将式(17)代入式(16),近似后验PDFs $q^{(j+1)}(y_t)$ 和 $q^{(j+1)}(r_t)$ 可以迭代更新为

$$\begin{cases} q^{(j+1)}(y_t) = (\hat{\varphi}_t)^{y_t}(1-\hat{\varphi}_t)^{1-y_t} \\ q^{(j+1)}(r_t) = (\hat{\zeta}_t)^{r_t}(1-\hat{\zeta}_t)^{1-r_t} \end{cases} \quad (27)$$

其中 $\hat{\varphi}_t, \hat{\zeta}_t$ 分别为估计的切换概率,可以表示为

$$\begin{cases} \hat{\varphi}_t = \frac{\varphi_{1,t}^{(j+1)}}{\varphi_{1,t}^{(j+1)} + \varphi_{2,t}^{(j+1)}} \\ \hat{\zeta}_t = \frac{\vartheta_{1,t}^{(j+1)}}{\vartheta_{1,t}^{(j+1)} + \vartheta_{2,t}^{(j+1)}} \end{cases}$$

中间参数分别由以下方程计算:

$$\begin{cases} \varphi_{1,t}^{(j+1)} = \exp\{E^{(j)}[\log \delta_t] - 0.5\text{tr}(\Psi_t^{(j+1)}E^{(j)}[\Sigma_t^{-1}])\} \\ \varphi_{2,t}^{(j+1)} = \exp\{E^{(j)}[\log(1-\delta_t)] - 0.5n_xE^{(j+1)}[\log \sigma_t] - \\ \quad 0.5E^{(j+1)}[\sigma_t]\text{tr}(\Psi_t^{(j+1)}E^{(j)}[\Sigma_t^{-1}])\} \\ \vartheta_{1,t}^{(j+1)} = \exp\{E^{(j)}[\log \lambda_t] - 0.5\text{tr}(\Xi_t^{(j+1)}E^{(j)}[R_t^{-1}])\} \\ \vartheta_{2,t}^{(j+1)} = \exp\{E^{(j)}[\log(1-\lambda_t)] - 0.5n_zE^{(j+1)}[\log \pi_t] - \\ \quad 0.5E^{(j+1)}[\pi_t]\text{tr}(\Xi_t^{(j+1)}E^{(j)}[R_t^{-1}])\} \end{cases} \quad (28)$$

**推导3** 将式(16)代入式(18)有

$$\begin{cases} \log q^{(j+1)}(y_t) = y_t\{E^{(j)}[\log \delta_t] - 0.5\text{tr}(\Psi_t^{(j+1)}E^{(j)}[\Sigma_t^{-1}])\} + \\ \quad (1-y_t)\{E^{(j)}[\log(1-\delta_t)] - 0.5n_xE^{(j+1)}[\log \sigma_t] - \\ \quad 0.5E^{(j+1)}[\sigma_t]\text{tr}(\Psi_t^{(j+1)}E^{(j)}[\Sigma_t^{-1}])\} + c_{y_t} = \\ \quad y_t\log \varphi_{1,t}^{(j+1)} + (1-y_t)\log \varphi_{2,t}^{(j+1)} + c_{y_t} \\ \log q^{(j+1)}(r_t) = r_t\{E^{(j)}[\log \lambda_t] - 0.5\text{tr}(\Xi_t^{(j+1)}E^{(j)}[R_t^{-1}])\} + \\ \quad (1-r_t)\{E^{(j)}[\log(1-\lambda_t)] - 0.5n_zE^{(j+1)}[\log \pi_t] - \\ \quad 0.5E^{(j+1)}[\pi_t]\text{tr}(\Xi_t^{(j+1)}E^{(j)}[R_t^{-1}])\} + c_{r_t} = \\ \quad r_t\log \vartheta_{1,t}^{(j+1)} + (1-r_t)\log \vartheta_{2,t}^{(j+1)} + c_{r_t} \end{cases} \quad (29)$$

对式(29)两边进行对数操作,并根据伯努利分

布的定义可以得到命题3中的式(27)、(28)。

**命题4** 该命题是在状态向量、噪声尺度参数等未知参数与噪声协方差矩阵 $\{\mathbf{x}_t, \sigma_t, \pi_t, y_t, r_t, \epsilon_t, \eta_t, \theta_t, \tau_t, \Sigma_t, R_t\}$ 保持不变的情况下实现切换概率 $\delta_t$ 和 $\lambda_t$ 的变分近似估计,分别使 $\Delta = \delta_t$ 和 $\Delta = \lambda_t$ 并将式(17)代入式(16),然后近似后验PDFs $q^{(j+1)}(\delta_t)$ 和 $q^{(j+1)}(\lambda_t)$ 可以被迭代更新为以下贝塔分布:

$$\begin{cases} q^{(j+1)}(\delta_t) = B(\delta_t; \hat{k}_{1,t}^{(j+1)}, \hat{k}_{2,t}^{(j+1)}) \\ q^{(j+1)}(\lambda_t) = B(\lambda_t; \hat{h}_{1,t}^{(j+1)}, \hat{h}_{2,t}^{(j+1)}) \end{cases} \quad (30)$$

其中形状参数 $\hat{k}_{1,t}^{(j+1)}, \hat{k}_{2,t}^{(j+1)}, \hat{h}_{1,t}^{(j+1)}, \hat{h}_{2,t}^{(j+1)}$ 分别为

$$\begin{cases} \hat{k}_{1,t}^{(j+1)} = k_0 + E^{(j+1)}[y_t] \\ \hat{k}_{2,t}^{(j+1)} = 2 - E^{(j+1)}[y_t] - k_0 \\ \hat{h}_{1,t}^{(j+1)} = h_0 + E^{(j+1)}[r_t] \\ \hat{h}_{2,t}^{(j+1)} = 2 - E^{(j+1)}[r_t] - h_0 \end{cases}$$

**推导4** 将式(16)代入式(18)有

$$\begin{aligned} \log q^{(j+1)}(\delta_t) &= (k_0 + E^{(j+1)}[y_t] - 1)\log \delta_t + \\ &\quad (1 - E^{(j+1)}[y_t] - k_0)\log(1 - \delta_t) + c_{\delta_t} = \\ &\quad (\hat{k}_{1,t}^{(j+1)} - 1)\log \delta_t + (\hat{k}_{2,t}^{(j+1)} - 1)\log(1 - \delta_t) + \\ &\quad c_{\delta_t} \log q^{(j+1)}(\lambda_t) = (h_0 + E^{(j+1)}[r_t] - 1)\log \lambda_t + \\ &\quad (1 - E^{(j+1)}[r_t] - h_0)\log(1 - \lambda_t) + c_{\lambda_t} = \\ &\quad (\hat{h}_{1,t}^{(j+1)} - 1)\log \lambda_t + (\hat{h}_{2,t}^{(j+1)} - 1)\log(1 - \lambda_t) + c_{\lambda_t} \end{aligned} \quad (31)$$

对式(31)的两边进行对数操作,并根据贝塔分布的定义可以得到命题4中的式(30)。

**命题5** 该命题是在状态向量、噪声尺度参数等未知参数与噪声协方差矩阵 $\{\mathbf{x}_t, \sigma_t, \pi_t, y_t, r_t, \delta_t, \lambda_t, \theta_t, \tau_t, \Sigma_t, R_t\}$ 保持不变的情况下实现分类分布参数 $\epsilon_t$ 和 $\eta_t$ 的变分近似估计,分别使 $\Delta = \epsilon_t$ 和 $\Delta = \eta_t$ 并将式(17)代入式(16),然后近似后验PDFs $q^{(j+1)}(\epsilon_t)$ 与 $q^{(j+1)}(\eta_t)$ 可以被迭代更新为以下分类分布:

$$\begin{cases} q^{(j+1)}(\epsilon_t) = \text{Cat}(\epsilon_t; \hat{\omega}_t^{(j+1)}, I) \\ q^{(j+1)}(\eta_t) = \text{Cat}(\eta_t; \hat{\varpi}_t^{(j+1)}, L) \end{cases} \quad (32)$$

其中混合概率向量 $\hat{\omega}_t^{(j+1)}, \hat{\varpi}_t^{(j+1)}$ 分别为

$$\begin{cases} \hat{\omega}_t^{(j+1)} = \omega_t^{(j+1)} / \sum_{i=1}^I \omega_{t,i}^{(j+1)} \\ \hat{\varpi}_t^{(j+1)} = \varpi_t^{(j+1)} / \sum_{l=1}^L \varpi_{t,l}^{(j+1)} \end{cases}$$

辅助参数向量 $\omega_{t,i}^{(j+1)}$ 与 $\varpi_{t,l}^{(j+1)}$ 分别为

$$\begin{cases} \omega_{t,i}^{(j+1)} = \exp\{(a_{0,i} - 1)E^{(j+1)}[\log \sigma_t] - \\ \quad b_{0,i}E^{(j+1)}[\sigma_t] + E^{(j)}[\log \theta_{t,i}]\} \\ \varpi_{t,l}^{(j+1)} = \exp\{(c_{0,l} - 1)E^{(j+1)}[\log \pi_t] - \\ \quad d_{0,l}E^{(j+1)}[\pi_t] + E^{(j)}[\log \tau_{t,l}]\} \end{cases} \quad (33)$$

**推导 5** 将式(16)代入式(18)有

$$\left\{ \begin{aligned} \log q^{(j+1)}(\boldsymbol{\varepsilon}_t) &= \sum_{i=1}^I \boldsymbol{\varepsilon}_{t,i} \exp\{(a_{0,i} - 1) \times E^{(j+1)}[\log \sigma_t] - \\ &\quad b_{0,i} E^{(j+1)}[\sigma_t] + E^{(j)}[\log \boldsymbol{\theta}_{t,i}]\} + \\ c_{\boldsymbol{\varepsilon}_t} &= \sum_{i=1}^I \boldsymbol{\varepsilon}_{t,i} \boldsymbol{\omega}_{t,i}^{(j+1)} + c_{\boldsymbol{\varepsilon}_t} \\ \log q^{(j+1)}(\boldsymbol{\eta}_t) &= \sum_{l=1}^L \boldsymbol{\eta}_{t,l} \exp\{(c_{0,l} - 1) \times E^{(j+1)}[\log \pi_t] - \\ &\quad d_{0,l} E^{(j+1)}[\pi_t] + E^{(j)}[\log \boldsymbol{\tau}_{t,l}]\} + \\ c_{\boldsymbol{\eta}_t} &= \sum_{l=1}^L \boldsymbol{\eta}_{t,l} \boldsymbol{\varpi}_{t,l}^{(j+1)} + c_{\boldsymbol{\eta}_t} \end{aligned} \right. \quad (34)$$

对式(34)的两边进行对数操作,并根据分类分布的定义可以得到命题 5 中的式(32)、(33)。

**命题 6** 该命题是在状态向量、噪声尺度参数等未知参数与噪声协方差矩阵  $\boldsymbol{\Theta} = \{\boldsymbol{x}_t, \sigma_t, \pi_t, y_t, r_t, \delta_t, \lambda_t, \boldsymbol{\varepsilon}_t, \boldsymbol{\eta}_t, \boldsymbol{\Sigma}_t, \boldsymbol{R}_t\}$  保持不变的情况下实现混合概率向量参数  $\boldsymbol{\theta}_t$  和  $\boldsymbol{\tau}_t$  的变分近似估计,分别使  $\Delta = \boldsymbol{\theta}_t$  和  $\Delta = \boldsymbol{\tau}_t$  并将式(17)代入式(16),然后近似后验 PDFs  $q^{(j+1)}(\boldsymbol{\theta}_t)$  和  $q^{(j+1)}(\boldsymbol{\tau}_t)$  可以迭代更新为狄利克雷分布:

$$\left\{ \begin{aligned} q^{(j+1)}(\boldsymbol{\theta}_t) &= \text{Dir}(\boldsymbol{\theta}_t; \hat{\boldsymbol{e}}_{t,t}^{(j+1)}, I) \\ q^{(j+1)}(\boldsymbol{\tau}_t) &= \text{Dir}(\boldsymbol{\tau}_t; \hat{\boldsymbol{g}}_{t,t}^{(j+1)}, L) \end{aligned} \right. \quad (35)$$

其中后验集中参数向量  $\hat{\boldsymbol{e}}_{t,t}^{(j+1)}$ 、 $\hat{\boldsymbol{g}}_{t,t}^{(j+1)}$  分别为

$$\left\{ \begin{aligned} \hat{\boldsymbol{e}}_{t,t}^{(j+1)} &= \boldsymbol{e}_{t,t-1} + E^{(j+1)}[\boldsymbol{\varepsilon}_t] \\ \hat{\boldsymbol{g}}_{t,t}^{(j+1)} &= \boldsymbol{g}_{t,t-1} + E^{(j+1)}[\boldsymbol{\eta}_t] \end{aligned} \right.$$

**推导 6** 将式(16)代入式(18)有

$$\left\{ \begin{aligned} \log q^{(j+1)}(\boldsymbol{\theta}_t) &= \sum_{i=1}^I (\hat{\boldsymbol{e}}_{t,t-1,i} + E^{(j+1)}[\boldsymbol{\varepsilon}_{t,i}] - 1) \log \boldsymbol{\theta}_{t,i} + c_{\boldsymbol{\theta}_t} \\ \log q^{(j+1)}(\boldsymbol{\tau}_t) &= \sum_{l=1}^L (\hat{\boldsymbol{g}}_{t,t-1,l} + E^{(j+1)}[\boldsymbol{\eta}_{t,l}] - 1) \log \boldsymbol{\tau}_{t,l} + c_{\boldsymbol{\tau}_t} \end{aligned} \right. \quad (36)$$

对式(36)两边进行对数操作,并根据狄利克雷分布的定义可以得到命题 6 中的式(35)。

**命题 7** 该命题是在状态向量、噪声尺度参数等未知参数  $\{\boldsymbol{x}_t, \sigma_t, \pi_t, y_t, r_t, \delta_t, \lambda_t, \boldsymbol{\varepsilon}_t, \boldsymbol{\eta}_t, \boldsymbol{\theta}_t, \boldsymbol{\tau}_t\}$  保持不变的情况下实现噪声协方差矩阵  $\boldsymbol{\Sigma}_t$  和  $\boldsymbol{R}_t$  的变分近似估计,分别使  $\Delta = \boldsymbol{\Sigma}_t$  和  $\Delta = \boldsymbol{R}_t$  并将式(17)代入式(16),然后近似后验 PDFs  $q^{(j+1)}(\boldsymbol{\Sigma}_t)$  和  $q^{(j+1)}(\boldsymbol{R}_t)$  可以迭代更新为逆威沙特分布:

$$\left\{ \begin{aligned} q^{(j+1)}(\boldsymbol{\Sigma}_t) &= \text{IW}(\boldsymbol{\Sigma}_t; \hat{\boldsymbol{s}}_{t,t}^{(j+1)}, \hat{\boldsymbol{S}}_{t,t}^{(j+1)}) \\ q^{(j+1)}(\boldsymbol{R}_t) &= \text{IW}(\boldsymbol{R}_t; \hat{\boldsymbol{u}}_{t,t}^{(j+1)}, \hat{\boldsymbol{U}}_{t,t}^{(j+1)}) \end{aligned} \right. \quad (37)$$

其中:后验自由度参数  $\hat{\boldsymbol{s}}_{t,t}^{(j+1)}$ 、 $\hat{\boldsymbol{u}}_{t,t}^{(j+1)}$  以及后验逆尺度矩阵  $\hat{\boldsymbol{S}}_{t,t}^{(j+1)}$ 、 $\hat{\boldsymbol{U}}_{t,t}^{(j+1)}$  分别为

$$\left\{ \begin{aligned} \hat{\boldsymbol{s}}_{t,t}^{(j+1)} &= \boldsymbol{s}_{t,t-1} + 1 \\ \hat{\boldsymbol{u}}_{t,t}^{(j+1)} &= \boldsymbol{u}_{t,t-1} + 1 \\ \hat{\boldsymbol{S}}_{t,t}^{(j+1)} &= \boldsymbol{S}_{t,t-1} + \{(1 - E^{(j+1)}[y_t])E^{(j+1)}[\sigma_t] + \\ &\quad E^{(j+1)}[y_t]\} \boldsymbol{\Psi}_t^{(j+1)} \\ \hat{\boldsymbol{U}}_{t,t}^{(j+1)} &= \boldsymbol{U}_{t,t-1} + \{(1 - E^{(j+1)}[r_t])E^{(j+1)}[\pi_t] + \\ &\quad E^{(j+1)}[r_t]\} \boldsymbol{\Xi}_t^{(j+1)} \end{aligned} \right.$$

**推导 7** 将式(16)代入式(18)有

$$\left\{ \begin{aligned} \log q^{(j+1)}(\boldsymbol{\Sigma}_t) &= -0.5(\boldsymbol{s}_{t,t-1} + \boldsymbol{n}_x + 2) \log |\boldsymbol{\Sigma}_t| - \\ &\quad 0.5 \text{tr}((\boldsymbol{S}_{t,t-1} + \{(1 - E^{(j+1)}[y_t])E^{(j+1)}[\sigma_t] + \\ &\quad E^{(j+1)}[y_t]\} \boldsymbol{\Psi}_t^{(j+1)}) \boldsymbol{\Sigma}_t^{-1}) + c_{\boldsymbol{\Sigma}_t} = -0.5(\hat{\boldsymbol{s}}_{t,t}^{(j+1)} + \\ &\quad \boldsymbol{n}_x + 1) \log |\boldsymbol{\Sigma}_t| - 0.5 \text{tr}(\hat{\boldsymbol{S}}_{t,t}^{(j+1)} \boldsymbol{\Sigma}_t^{-1}) + c_{\boldsymbol{\Sigma}_t} \\ \log q^{(j+1)}(\boldsymbol{R}_t) &= -0.5(\boldsymbol{u}_{t,t-1} + \boldsymbol{n}_z + 2) \log |\boldsymbol{R}_t| - \\ &\quad 0.5 \text{tr}((\boldsymbol{U}_{t,t-1} + \{(1 - E^{(j+1)}[r_t])E^{(j+1)}[\pi_t] + \\ &\quad E^{(j+1)}[r_t]\} \boldsymbol{\Xi}_t^{(j+1)}) \boldsymbol{R}_t^{-1}) + c_{\boldsymbol{R}_t} = -0.5(\hat{\boldsymbol{u}}_{t,t}^{(j+1)} + \\ &\quad \boldsymbol{n}_z + 1) \log |\boldsymbol{R}_t| - 0.5 \text{tr}(\hat{\boldsymbol{U}}_{t,t}^{(j+1)} \boldsymbol{R}_t^{-1}) + c_{\boldsymbol{R}_t} \end{aligned} \right. \quad (38)$$

对式(38)的两边进行对数操作,并根据逆威沙特分布的定义可以得到命题 7 中式(37)。

在式(19)~式(37)所示的近似后验 PDF 的迭代更新中,涉及的模型参数期望可计算为

$$\left\{ \begin{aligned} E^{(j+1)}[\sigma_t] &= \hat{\alpha}_{1,t}^{(j+1)} / \hat{\alpha}_{2,t}^{(j+1)} \\ E^{(j+1)}[\pi_t] &= \hat{\beta}_{1,t}^{(j+1)} / \hat{\beta}_{2,t}^{(j+1)} \\ E^{(j+1)}[\log \sigma_t] &= \psi(\hat{\alpha}_{1,t}^{(j+1)}) - \log(\hat{\alpha}_{2,t}^{(j+1)}) \\ E^{(i+1)}[y_t] &= \hat{\varphi}_t \\ E^{(i+1)}[r_t] &= \hat{\gamma}_t \\ E^{(j+1)}[\log \pi_t] &= \psi(\hat{\beta}_{1,t}^{(j+1)}) - \log(\hat{\beta}_{2,t}^{(j+1)}) \\ E^{(j+1)}[\log \delta_t] &= \psi(\hat{r}_{1,t}^{(j+1)}) - \psi(2) \\ E^{(j+1)}[\log \lambda_t] &= \psi(\hat{h}_{1,t}^{(j+1)}) - \psi(2) \\ E^{(j+1)}[\log \boldsymbol{\theta}_{t,i}] &= \psi(\hat{n}_{t,i}^{(j+1)}) - \psi(\sum_{i=1}^I \hat{n}_{t,i}^{(j+1)}) \\ E^{(j+1)}[\log \boldsymbol{\tau}_{t,l}] &= \psi(\hat{\boldsymbol{g}}_{t,l}^{(j+1)}) - \psi(\sum_{l=1}^L \hat{\boldsymbol{g}}_{t,l}^{(j+1)}) \\ E^{(j+1)}[\boldsymbol{\Sigma}_t^{-1}] &= \hat{\boldsymbol{s}}_{t,t}^{(j+1)} (\boldsymbol{S}_{t,t}^{(j+1)})^{-1} \\ E^{(j+1)}[\boldsymbol{R}_t^{-1}] &= \hat{\boldsymbol{u}}_{t,t}^{(j+1)} (\hat{\boldsymbol{U}}_{t,t}^{(j+1)})^{-1} \\ E^{(j+1)}[\boldsymbol{\varepsilon}_t] &= \hat{\boldsymbol{\omega}}_t^{(j+1)} \\ E^{(j+1)}[\boldsymbol{\eta}_t] &= \hat{\boldsymbol{\varpi}}_t^{(j+1)} \end{aligned} \right. \quad (39)$$

根据式(12)~式(14)中的时间更新与式(19)~式(39)所示的量测更新给出了如下的 GHTSRKF 的一次迭代的执行过程,其中算法输入为上一时刻的

后验估计结果。

**注4** 在提出的 GHTSRKF 中,其一步状态预测 PDF 和量测似然 PDF 分别建模为式(11)中 GHTS 分布,通过标准高斯分布与重尾分布之间的自适应切换以适应非平稳重尾状态与量测噪声。然后对于两种 PDFs 中的状态向量、噪声协方差矩阵、噪声尺度参数、伯努利参数等未知参数根据 VB 理论实现式(16)中变分近似并利用式(19)~式(39)中固定点迭代方法进行近似估计。最后得到 GHTSRKF 算法如下,其通过时间更新与迭代量测更新过程实现了非高斯噪声条件下的状态估计。

**算法:**提出的 GHTSRKF

**输入:**  $\hat{\mathbf{x}}_{t-1|t-1}$ ,  $\mathbf{P}_{t-1|t-1}$ ,  $\mathbf{z}_t$ ,  $\hat{\mathbf{u}}_{t-1|t-1}$ ,

$\hat{\mathbf{U}}_{t-1|t-1}$ ,  $\hat{\mathbf{e}}_{t-1|t-1}$ ,  $\hat{\mathbf{g}}_{t-1|t-1}$

**时间更新:**

通过式(12)~式(14)计算先验参数。

**量测更新:**

利用式(39)计算初始参数期望。

for  $j=0:J-1$

由式(21)计算  $\hat{\mathbf{P}}_{t|t-1}^{(j+1)}$  和  $\hat{\mathbf{R}}_t^{(j+1)}$ ;

根据式(19)更新  $q^{(j+1)}(\mathbf{x}_t)$ ;

由式(26)~式(39)更新  $q^{(j+1)}(\boldsymbol{\Sigma}_t)$ 、 $q^{(j+1)}(\mathbf{R}_t)$ ;

end

$\hat{\mathbf{x}}_{t|t} = \hat{\mathbf{x}}_{t|t}^{(J)}$ ,  $\mathbf{P}_{t|t} = \mathbf{P}_{t|t}^{(J)}$ ,  $\hat{\mathbf{u}}_{t|t} = \hat{\mathbf{u}}_{t|t}^{(J)}$ ,

$\hat{\mathbf{U}}_{t|t} = \hat{\mathbf{U}}_{t|t}^{(J)}$ ,  $\hat{\mathbf{e}}_{t|t} = \hat{\mathbf{e}}_{t|t}^{(J)}$ ,  $\hat{\mathbf{g}}_{t|t} = \hat{\mathbf{g}}_{t|t}^{(J)}$

**输出:**  $\hat{\mathbf{x}}_{t|t}$ ,  $\mathbf{P}_{t|t}$ ,  $\hat{\mathbf{u}}_{t|t}$ ,  $\hat{\mathbf{U}}_{t|t}$ ,  $\hat{\mathbf{e}}_{t|t}$ ,  $\hat{\mathbf{g}}_{t|t}$

### 3.3 对提出的 GHTSRKF 的讨论

首先分析了 GHTSRKF 与现有的滤波器之间的关联。当混合伽马分布的数量  $I=L=1$  时,混合伽马分布会退化为单一伽马分布,这时 GHTSRKF 几乎等同于 M2GRKF<sup>[23]</sup> 和噪声尺度矩阵未知情况下的 GSTMRKF<sup>[20]</sup> 的滤波框架。与此同时,当先验形状参数  $k_0=h_0=1$  时,GHTSRKF 会进一步退化为变分自适应 KF<sup>[25]</sup>;当  $k_0=h_0=0$  时,GHTS-RKF 基本与现有的 ORSTKF<sup>[18]</sup> 的形式一样。从上述分析可知,变分自适应 KF, ORSTKF, GST-MRKF 和 M2GRKF 是本文提出的滤波器的几种特殊形式,这在一定程度上说明了所提出的滤波器性能更全面。

在文献[22]的 GSTM 分布和文献[23]的 M2G 分布中,单一伽马分布被用来建模尺度参数以近似重尾噪声的特性,但是动态野值特点是很复杂的,单一伽马分布所包含先验信息在这样的情况下过于简单而不能准确的近似重尾特性。而在 GHTS 分布中,本文构造的混合伽马分布可以提供一个粗略先

验信息集合并通过自适应学习混合概率向量从该集合中提取一个准确先验用于建模尺度参数。因此理论上,如此的 GHTS 分布可以更准确描述非平稳重尾噪声,基于该分布的 GHTSRKF 的鲁棒性能将会更优越。

在一般的应用中,考虑到大概率情况下噪声是不准确的高斯噪声,因此有  $0 < 1 - \delta_t \ll \delta_t < 1$  和  $0 < 1 - \lambda_t \ll \lambda_t < 1$  且  $E[\delta_t] = k_0$  和  $E[\lambda_t] = h_0$ ,所以在本文中设先验形状参数  $k_0 = h_0 = 0.85$ 。对于尺度参数  $0 < \sigma_t \leq 1$  和  $0 < \pi_t \leq 1$ ,根据式(11)可以得到

$$p(\sigma_t) = \sum_{i=1}^I \theta_{t,i} \sigma_{t,i} \text{ 和 } p(\pi_t) = \sum_{l=1}^L \tau_{t,l} \pi_{t,l}, \text{ 且各分量的期望为 } E[\sigma_{t,i}] = a_{0,i}/b_{0,i} \text{ 和 } E[\pi_{t,l}] = c_{0,l}/d_{0,l}。 \text{ 本文设初始形状参数向量 } \mathbf{a}_0 = \mathbf{c}_0 = \mathbf{2}_{1 \times L}, \text{ 初始速率参数向量 } \mathbf{b}_0 = \mathbf{d}_0 = [10, 100, 1\,000, \dots, 10^L]_{1 \times L}。$$

$$\begin{cases} \hat{\mathbf{P}}_{t|t-1}^{(j+1)} = \frac{m_t \tilde{\mathbf{P}}_{t|t-1} / \Phi_t^{(j+1)} + \Psi_t^{(j+1)}}{m_t + 1} \\ \hat{\mathbf{R}}_t^{(j+1)} = \frac{\rho \hat{\mathbf{U}}_{t-1|t-1} \hat{\mathbf{R}}_{t-1} / \Omega_t^{(j+1)} + \Xi_k^{(j+1)}}{\rho \hat{\mathbf{u}}_{t-1|t-1} + 1} \end{cases} \quad (40)$$

式中:  $\tilde{\mathbf{P}}_{t|t-1}$  由式(14)得到,  $\hat{\mathbf{R}}_{t-1}$  为  $t-1$  时刻的估计量测噪声协方差矩阵。

辅助参数  $\Phi_t^{(j+1)}$ 、 $\Omega_t^{(j+1)}$  分别为

$$\begin{cases} \Phi_t^{(j+1)} = (1 - E^{(j+1)}[y_t]) E^{(j+1)}[\sigma_t] + E^{(j+1)}[y_t] \\ \Omega_t^{(j+1)} = (1 - E^{(j+1)}[r_t]) E^{(j+1)}[\pi_t] + E^{(j+1)}[r_t] \end{cases} \quad (41)$$

由式(40)可知,调节参数  $m_t$  可视为一个用来平衡  $\Psi_t^{(j+1)}$  与自适应校正项  $\tilde{\mathbf{P}}_{t|t-1} / \Phi_t^{(j+1)}$  的调和权重系数。调节参数  $m_t$  越大,由粗略预选的系统噪声协方差引起的先验不确定性就越多;相反当  $m_t$  较小时,此协方差中固有的经验信息就容易被忽略,该参数的影响在仿真中会被分析。同样,遗忘因子  $\rho$  也是为了平衡  $\hat{\mathbf{R}}_{t-1} / \Omega_t^{(j+1)}$  和  $\Xi_k^{(j+1)}$ ,考虑到量测噪声协方差是缓慢变化的,设  $\rho = 1 - \exp(-4)$ <sup>[25]</sup>。

对于固定点迭代数量  $J$  的选取需要平衡算法的估计精度与计算复杂度,该参数对滤波器性能的影响在仿真中会被讨论。

对于一个动态系统来说,其系统不确定性主要有两个方面:噪声不确定性(系统与量测噪声)与系统参数不确定性(未知输入、失配的系统与量测矩阵)。本文提出的鲁棒滤波器主要是为降低系统与量测噪声的不确定所带来的负面影响,没有考虑系统参数的不确定所带的影响。受限于时间与现有的

知识储备,在本文中并没有针对系统参数的不确定问题进行研究,在下一步的研究中,结合本文方法对系统参数的不确定问题进行深入研究。

## 4 仿真与分析

### 4.1 设置

本文使用文献[23]中的仿真实例来验证所提出的 GHTSRKF 的性能。该模型采用类似式(1)的状态空间模型来描述,状态为  $\mathbf{x}_t = [x_t, y_t, \dot{x}_t, \dot{y}_t]$  其中  $x_t$  和  $\dot{x}_t, y_t$  和  $\dot{y}_t$  分别为位置和对应的速度。状态转移矩阵  $\mathbf{F}_t = \begin{bmatrix} \mathbf{I}_2 & \Delta t \mathbf{I}_2 \\ \mathbf{0}_{2 \times 2} & \mathbf{I}_2 \end{bmatrix}$  和量测矩阵  $\mathbf{H}_t = [\mathbf{I}_2 \quad \mathbf{0}_{2 \times 2}]$ , 其中  $\Delta t = 1$  s,  $\mathbf{I}_2$  是 2 维单位矩阵。

考虑以下场景验证所提滤波器的性能。

**场景 1** 时变高斯噪声为:

$$\mathbf{w}_{t-1} \sim N(0, \mathbf{Q}_{t-1}), \mathbf{v}_t \sim N(0, \mathbf{R}_t) \quad (42)$$

**场景 2** 非平稳重尾噪声为

$$\begin{cases} \mathbf{w}_{t-1} \sim q_1 N(0, \mathbf{Q}_{t-1}) + (1-q_1) N(0, 100 \mathbf{Q}_{t-1}), & t \in [201, 400] \\ \mathbf{w}_{t-1} \sim q_2 N(0, \mathbf{Q}_{t-1}) + (1-q_2) N(0, 100 \mathbf{Q}_{t-1}), & t \in [601, 800] \\ \mathbf{w}_{t-1} \sim N(0, \mathbf{Q}_{t-1}), & \text{其他时间} \\ \mathbf{v}_t \sim q_1 N(0, \mathbf{R}_t) + (1-q_1) N(0, 200 \mathbf{R}_t), & t \in [201, 400] \\ \mathbf{v}_t \sim q_2 N(0, \mathbf{R}_t) + (1-q_2) N(0, 200 \mathbf{R}_t), & t \in [601, 800] \\ \mathbf{v}_t \sim N(0, \mathbf{R}_t), & \text{其他时间} \end{cases} \quad (43)$$

式中:  $q_1 = 0.85, q_2 = 0.95$  分别表示概率,真实的时变噪声协方差矩阵  $\mathbf{Q}_{t-1}, \mathbf{R}_t$  分别为

$$\begin{cases} \mathbf{Q}_{t-1} = [6.5 + 0.5 \cos(\pi t/T)] \times \\ \quad \begin{bmatrix} (\Delta t^3/3) \mathbf{I}_2 & (\Delta t^2/2) \mathbf{I}_2 \\ (\Delta t^2/2) \mathbf{I}_2 & \Delta t \mathbf{I}_2 \end{bmatrix} \\ \mathbf{R}_t = [0.1 + 0.05 \cos(\pi t/T)] \times \\ \quad 100 \begin{bmatrix} \Delta t & (\Delta t^2/2) \\ (\Delta t^2/2) & \Delta t \end{bmatrix} \end{cases} \quad (44)$$

仿真参数为  $\hat{\mathbf{x}}_0 = [0, 0, 10, 10]^T, \tilde{\mathbf{Q}}_0 = \mathbf{I}_4, \tilde{\mathbf{R}}_0 = 100 \mathbf{I}_2, \mathbf{P}_0 = 100 \mathbf{I}_4, T = 1\ 000$ 。而且,本文将有不准确的噪声协方差矩阵的 KF(KFNNC),有准确的噪声协方差矩阵的 KF(KFTNC),现有的 ORSTKF<sup>[18]</sup>, HTMRKF<sup>[19]</sup>, GSTMRKF<sup>[22]</sup> 和 M2GRKF<sup>[23]</sup> 作为对比组进行测试。在提出的 GHTSRKF 中,滤波器参数设置为  $J=20, I=L=7, m_0=4, u_0=5, \mathbf{U}_0=\tilde{\mathbf{R}}_0$ 。而且对比方法 ORSTKF, HTMRKF, GSTMRKF 和 M2GRKF 中的参数设置分别与文献[18-21]中的设置保持一致。

为了从统计意义上评估不同滤波器的估计性

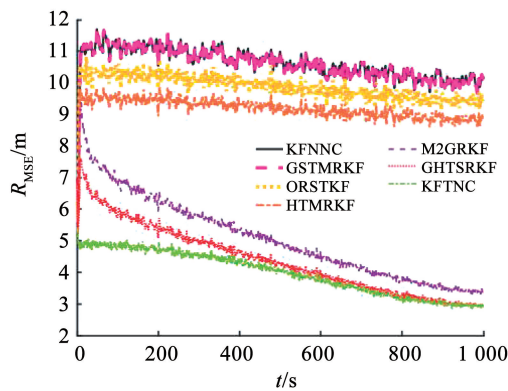
能,本文将均方根误差(root mean square error, RMSE)和平均均方根误差(average root mean square error, ARMSE)作为性能指标,分别定义为

$$\begin{cases} R_{\text{MSE}_{\text{pos}}} = \sqrt{\frac{\sum_{m=1}^m ((x_t^m - \hat{x}_t^m)^2 + (y_t^m - \hat{y}_t^m)^2)}{m}} \\ A_{\text{RMSE}_{\text{pos}}} = \sqrt{\frac{\sum_{t=1}^T \sum_{m=1}^m ((x_t^m - \hat{x}_t^m)^2 + (y_t^m - \hat{y}_t^m)^2)}{mT}} \end{cases} \quad (45)$$

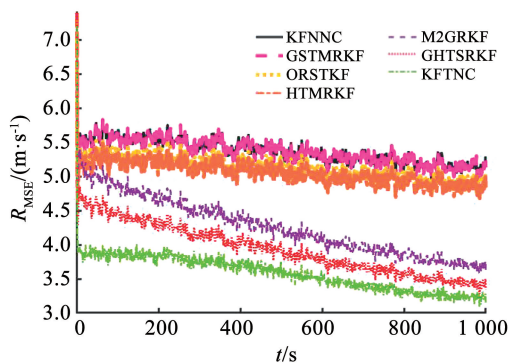
式中:  $R_{\text{MSE}_{\text{pos}}}, A_{\text{RMSE}_{\text{pos}}}$  分别为位置的均方根误差和平均均方根误差(其中  $R_{\text{MSE}}, A_{\text{RMSE}}$  分别为均方根误差和平均均方根误差, pos 为位置),  $(x_t^m, y_t^m), (\hat{x}_t^m, \hat{y}_t^m)$  分别为在第  $m$  次蒙特卡洛仿真时的真实位置与估计位置,蒙特卡洛仿真次数为 1 000;由位置的  $R_{\text{MSE}}$  和  $A_{\text{RMSE}}$  可以写出速度的性能指标表达式。

### 4.2 估计精度

首先,在不准确时变高斯噪声的情况下,分别来自现有的滤波器与提出的滤波器的位置与速度的估计结果在图 2 中呈现。



(a) 位置估计



(b) 速度估计

图 2 场景 1 中位置与速度的  $R_{\text{MSEs}}$

Fig. 2  $R_{\text{MSEs}}$  of position and velocity in Case 1

由图 2 可知,所提出的 GHTSRKF 的精度优于

其他 RKF<sub>s</sub>, 它的  $R_{MSEs}$  最接近噪声信息准确的 KFTNC 的  $R_{MSEs}$ 。由于不准确的时变噪声的影响, KFNNC 的精度严重退化。GSTMRKF 的估计结果与 KFNNC 类似, 虽然此滤波器设计时考虑了高斯噪声的情况, 但预选的噪声尺度矩阵不准确严重影响了 GSTMRKF 的性能。其次, 由于 ORSTKF 和 HTMRKF 只考虑了存在野值的情况, 因此它们估计效果不理想。M2GRKF 受益于两个高斯分布混合框架, 在噪声是未知时变高斯的情况下依然取得了较好的估计结果。然而由于单一的粗糙先验信息, 现有的 M2GRKF 与提出的 GHTSRKF 相比并不占优势。因此可以得到结论, 当系统与量测噪声是未知时变高斯噪声时, 相比于现有的滤波器, GHTSRKF 具有更好的估计性能。

然后, 非平稳重尾噪声情况下, 分别来自现有的滤波器与提出的滤波器的位置与速度  $R_{MSEs}$  被展示在图 3 中。

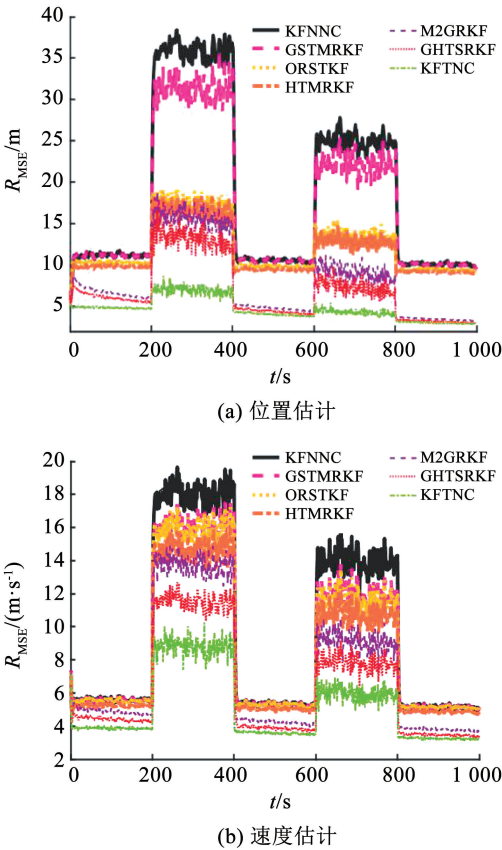


图3 场景2中位置与速度的  $R_{MSEs}$

Fig.3  $R_{MSEs}$  of position and velocity in Case 2

由图 3 可知, 提出的 GHTSRKF 在所有时段的位置与速度估计结果都明显优于现有的 RKF<sub>s</sub>, 这证明了 GHTS 分布更适合对非平稳重尾噪声建模。在 1~200、401~600、801~1 000 s 3 个时间段内, 各滤波器的表现与图 1 结果基本一致。在 201~400、601~800 s 时间段内不同的重尾噪声情况下,

KFNNC 的性能出现不同程度的恶化, 这也说明了发展与研究 RKF 的必要性。GSTMRKF 表现出了一定的鲁棒性, 但由于承受着噪声尺度矩阵需要事先已知的限制, 鲁棒性能依然较差。ORSTKF 和 HTMRKF 由于所构建的分层高斯状态模型而对这种情况有着不错的鲁棒性能, 并且 HTMRKF 受益于其采用的重尾混合分布在性能上以微弱优势赢得了 ORSTKF。由于针对非平稳重尾噪声的设计, 现有的 M2GRKF 呈现出了更好的估计性能, 然而 M2GRKF 在估计精度方面比提出的 GHTSRKF 表现略差。本文猜测应该是混合伽马分布能够预先提供丰富的粗略先验信息, 并通过混合概率向量的自适应学习提取出一个准确的先验, 这使得提出的滤波器能够对不同的重尾噪声进行更加准确描述, 因此鲁棒性也更强。综合所有时间段的估计结果, 可以发现提出的滤波器受益于高斯和新设计的重尾分布的切换和多重伽马分布混合对尺度参数的建模, 其性能在非平稳重尾噪声情况下明显优于现有的 RKF<sub>s</sub>。

除此之外, 本文比较了所提出的滤波器与其他滤波器算法的每步迭代计算时间, 结果见表 1。从表 1 可以看出, 在上述滤波器中, 标准 CKF 的计算速度最快。由于状态和未知参数的 VB 迭代极大地增加了算法的计算负荷, 各鲁棒滤波器算法每一步迭代时间明显增加。与 RSTKF、GSTMRKF、HTMRKF 和 M2GRKF 相比, 本文滤波器的运行时间略有增加。这说明本文滤波器在略微增加计算负担的基础上实现了更高的状态估计精度。

表1 各算法单次运行的时间

| Tab.1 Run times in a signal step from each algorithms |        |         |         |
|-------------------------------------------------------|--------|---------|---------|
| 算法                                                    | KF     | GSTMRKF | ORSTKF  |
| 时间/ms                                                 | 0.04   | 0.97    | 2.34    |
| 算法                                                    | HTMRKF | M2GRKF  | GHTSRKF |
| 时间/ms                                                 | 3.74   | 4.28    | 4.71    |

4.3 参数影响

本文分析了最大迭代数目  $J$ , 调节参数  $m_i$  和伽马分布的混合数量  $I$  和  $L$  对所提出的滤波器性能的影响, 在场景 2 下该滤波器的位置与速度的  $A_{RMSEs}$  分别展示在图 4~6 中。

图 4 显示了当迭代次数  $J \geq 6$  时, 滤波器收敛, 这说明本文的 GHTSRKF 具有满意的收敛速度。而且也说明了在满足估计精度的同时可以设定合理的固定点迭代数目以提升滤波器的计算效率。图 5 呈现了当固定点迭代次数  $J \in [1, 20]$ , 调节参数  $m_i \in [1, 20]$  时提出的滤波器的位置与速度估计。结果

显示,当  $m_i > 4$  时,调节参数对  $A_{\text{RMSEs}}$  的影响是有限的,  $m_i = 4$  时估计精度相对最优,从图 5 中可得该参数选择的区间范围推荐为  $m_i \in [3, 8]$ 。

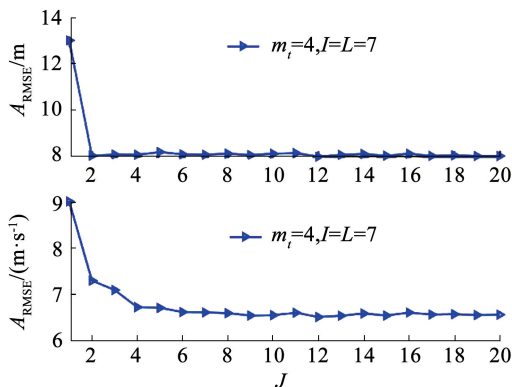


图 4  $J=1, 2, \dots, 20$  时 GHTSRKF 的  $A_{\text{RMSEs}}$

Fig. 4  $A_{\text{RMSEs}}$  of GHTSRKF with  $J=1, 2, \dots, 20$

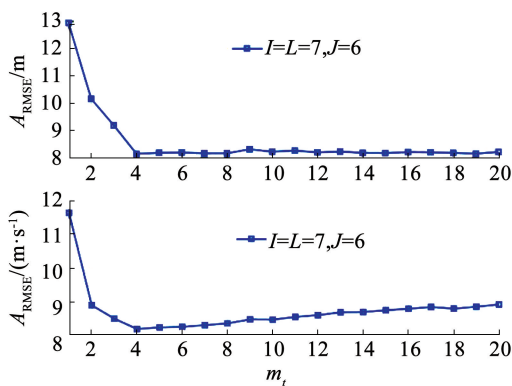


图 5  $m_i=1, 2, \dots, 20$  时 GHTSRKF 的  $A_{\text{RMSEs}}$

Fig. 5  $A_{\text{RMSEs}}$  of GHTSRKF with  $m_i=1, 2, \dots, 20$

图 6 呈现了当固定迭代次数  $J$  在区间  $[1, 20]$  时,伽马分布数量  $I=L \in [1, 20]$  时提出的滤波器的位置与速度估计,结果显示所提出的 GHTSRKF 的精度应随着伽马分布混合的数量  $I=L$  的增加而单调上升。然而,当  $I=L \geq 7$  时,滤波器的估计精度的提升很小,因此通过仔细选择  $I$  和  $L$  的大小可以平衡算法的估计精度和计算复杂度。

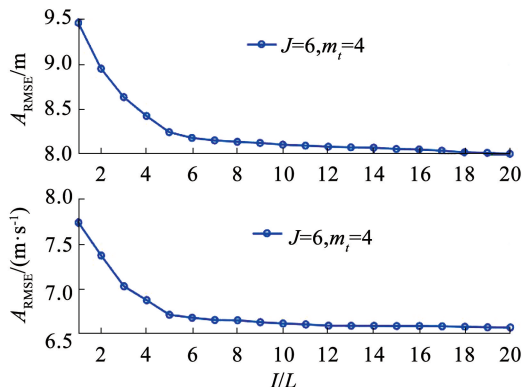


图 6  $I=L=1, 2, \dots, 20$  时 GHTSRKF 的  $A_{\text{RMSEs}}$

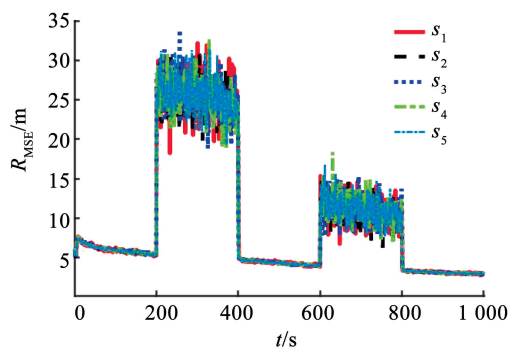
Fig. 6  $A_{\text{RMSEs}}$  of GHTSRKF with  $I=L=1, 2, \dots, 20$

#### 4.4 初始状态影响

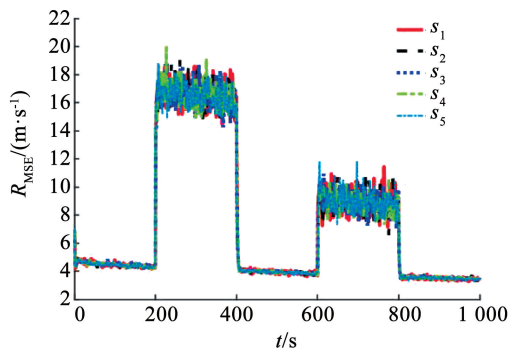
本文分析了不同的初始状态对所提出的滤波器性能的影响,设置 5 组不同的初始状态值,分别为

$$\begin{cases} s_1: \hat{\mathbf{x}}_0 = [0, 0, 0, 0]^T \\ s_2: \hat{\mathbf{x}}_0 = [100, 100, 10, 10]^T \\ s_3: \hat{\mathbf{x}}_0 = [500, 500, 100, 100]^T \\ s_4: \hat{\mathbf{x}}_0 = [-100, -100, 10, 10]^T \\ s_5: \hat{\mathbf{x}}_0 = [-500, -500, 100, 100]^T \end{cases} \quad (46)$$

在场景 2 下该滤波器的位置与速度的  $R_{\text{MSEs}}$  展示在图 7 中。



(a) 位置估计



(b) 速度估计

图 7 不同初始状态情况下 GHTSRKF 的估计

Fig. 7 Estimation of GHTSRKF with different initial states

由图 7 可知,所提出的 GHTSRKF 算法的估计精度对初始状态  $\hat{\mathbf{x}}_0$  的选取不敏感,其位置与速度的估计结果基本保持一致。因此可得初始状态对所提出算法的估计性能基本没有影响。

## 5 结 论

1) 所提出的 GHTSRKF 算法的估计精度对初始状态的选取不敏感,其位置与速度的估计结果基本保持一致,因此可得初始状态对所提出算法的估计性能基本没有影响。

2) 所提出的 GHTSRKF 的精度优于其他 RKF, 它的  $R_{\text{MSEs}}$  最接近噪声信息准确的 KFTNC 的  $R_{\text{MSEs}}$ 。

3) 当系统与量测噪声是未知时变高斯噪声时,

相比于现有的滤波器,GHTSRKF 具有更好的估计性能,且该滤波器能够对不同的重尾噪声进行更加准确描述,因此鲁棒性也更强。

参考文献

[1]SARGOLZAEI A, YAZDANI K, ABBASPOUR A, et al. Detection and mitigation of false data injection attacks in networked control systems[J]. IEEE Transactions on Industrial Informatics, 2020, 16(6): 4281. DOI: 10.1109/TII.2019.2952067

[2]LI Haoqing, MEDINA D, VILÀ-VALLS J, et al. Robust variational-based Kalman filter for outlier rejection with correlated measurements [J]. IEEE Transactions on Signal Processing, 2020, 69: 357. DOI: 10.1109/TSP.2020.3042944

[3]潘泉, 胡玉梅, 兰华, 等. 信息融合理论研究进展: 基于变分贝叶斯的联合优化[J]. 自动化学报, 2019, 45(7): 1207

PAN Quan, HU Yumei, LAN Hua, et al. Information fusion progress: Joint optimization based on variational Bayesian theory[J]. Acta Automatica Sinica, 2019, 45 ( 7 ): 1207. DOI: 10.16383/j. aas. c180029

[4]罗小元, 潘雪扬, 王新宇, 等. 基于自适应 Kalman 滤波的智能电网假数据注入攻击检测[J]. 自动化学报, 2022, 48(12): 2960

LUO Xiaoyuan, PAN Xueyang, WANG Xinyu, et al. Detection of false data injection attack in smart grid via adaptive Kalman filtering [J]. Acta Automatica Sinica, 2022(12): 2960. DOI: 10.16383/j. aas. c190636

[5]HU Yumei, WANG Xuezhi, PAN Quan, et al. Variational Bayesian Kalman filter using natural gradient [J]. Chinese Journal of Aeronautics, 2022, 35(5): 1. DOI: 10.1016/j. cja.2021.08.033

[6]LI Ling, SONG Xinmin. State estimation for systems with packet dropping and state equality constraints[J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2019, 66(9): 1572. DOI: 10.1109/TCSII.2018.2889047

[7]WANG Jian, ZHANG Tao, JIN Bonan, et al. Student's t-based robust Kalman filter for a SINS/USBL integration navigation strategy [J]. IEEE Sensors Journal, 2020, 20(10): 5540. DOI: 10.1109/JSEN.2020.2970766

[8]HUANG Yulong, ZHANG Yonggang, XU Bo, et al. A new adaptive extended Kalman filter for cooperative localization [J]. IEEE Transactions on Aerospace and Electronic Systems, 2018, 54(1): 353. DOI: 10.1109/TAES.2017.2756763

[9]ZHANG Junhao, LI Peng, JIN Congcong, et al. A novel adaptive Kalman filtering approach to human motion tracking with magnetic-inertial sensors[J]. IEEE Transactions on Industrial Electronics, 2020, 67(10): 8659. DOI: 10.1109/TIE.2019.2946557

[10]ZHU Hao, ZHANG Guorui, LI Yongfu, et al. An adaptive Kalman filter with inaccurate noise covariances in the presence of outliers [J]. IEEE Transactions on Automatic Control, 2022, 67(1): 374. DOI: 10.1109/TAC.2021.3056343

[11]WANG Guangcai, XU Xiaosu, ZHANG Tao. M-M estimation-based robust cubature Kalman filter for INS/GPS integrated navigation system [J]. IEEE Transactions on Instrumentation and Measurement, 2020, 70: 9501511. DOI: 10.1109/TIM.2020.3021224

[12]QIU Zhenbing, GUO Lei. Improved cubature Kalman filter for spacecraft attitude estimation [J]. IEEE Transactions on

Instrumentation and Measurement, 2021, 70: 9504213. DOI: 10.1109/TIM.2020.3041077

[13]FAN Xuxiang, WANG Gang, HAN Jiachen, et al. Interacting multiple model based on maximum correntropy Kalman filter[J]. IEEE Transactions on Circuits and Systems II: Express Briefs, 2021, 68(8): 3017. DOI: 10.1109/TCSII.2021.3068221

[14]CHEN Badong, LIU Xi, ZHAO Haiquan, et al. Maximum correntropy Kalman filter[J]. Automatica, 2017, 76: 70. DOI: 10.1016/j. automatica.2016.10.004

[15]CHEN Badong, DANG Lujuan, GU Yuantao, et al. Minimum error entropy Kalman filter[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, 51(9): 5819. DOI: 10.1109/TSMC.2019.2957269

[16]TRONARP F, KARVONEN T, SÄRKKÄ S. Student's t-filters for noise scale estimation[J]. IEEE Signal Processing Letters, 2019, 26(2): 352. DOI: 10.1109/LSP.2018.2889440

[17]HUANG Yulong, ZHANG Yonggang, LI Ning, et al. A novel robust student's t-based Kalman filter[J]. IEEE Transactions on Aerospace and Electronic Systems, 2017, 53(3): 1545. DOI: 10.1109/TAES.2017.2651684

[18]HUANG Yulong, ZHANG Yonggang, XU Bo, et al. A new outlier-robust student's t based Gaussian approximate filter for cooperative localization[J]. IEEE/ASME Transactions on Mechatronics, 2017, 22(5): 2380. DOI: 10.1109/TMECH.2017.2744651

[19]BAI Mingming, HUANG Yulong, ZHANG Yonggang, et al. A novel heavy-tailed mixture distribution based robust Kalman filter for cooperative localization [J]. IEEE Transactions on Industrial Informatics, 2021, 17 ( 5 ): 3671. DOI: 10.1109/TII.2020.3015001

[20]BAI Mingming, HUANG Yulong, CHEN Badong, et al. A novel mixture distributions-based robust Kalman filter for cooperative localization[J]. IEEE Sensors Journal, 2020, 20(24): 14994. DOI: 10.1109/JSEN.2020.3012153

[21]XU Bo, GUO Yu, HU Junmiao. An improved robust Kalman filter for SINS/DVL tightly integrated navigation system [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 8502915. DOI: 10.1109/TIM.2021.3079556

[22]HUANG Yulong, ZHANG Yonggang, ZHAO Yuxin, et al. A novel robust gaussian-student's t mixture distribution based Kalman filter [J]. IEEE Transactions on Signal Processing, 2019, 67(13): 3606. DOI: 10.1109/TSP.2019.2916755

[23]ZHU Hao, ZHANG Guorui, LI Yongfu, et al. A novel robust Kalman filter with unknown non-stationary heavy-tailed noise[J]. Automatica, 2021, 127: 109511. DOI: 10.1016/j. automatica.2021.109511

[24]ZHU Hao, LEUNG H, HE Zhongshi. State estimation in unknown non-Gaussian measurement noise using variational Bayesian technique[J]. IEEE Transactions on Aerospace and Electronic Systems, 2013, 49(4): 2601. DOI: 10.1109/TAES.2013.6621839

[25]HUANG Yulong, ZHANG Yonggang, WU Zheming, et al. A novel adaptive Kalman filter with inaccurate process and measurement noise covariance matrices[J]. IEEE Transactions on Automatic Control, 2018, 63(2): 594. DOI: 10.1109/TAC.2017.2730480