

# 非对称偏斜噪声条件下一种鲁棒概率系统辨识算法研究

刘鑫<sup>1</sup> 陈强<sup>2</sup> 王兰豪<sup>3</sup> 代伟<sup>1,2</sup>

**摘要** 在现有的系统辨识算法中,常用的高斯、学生氏  $t$  (Student's  $t$ , St)、拉普拉斯等噪声分布均呈现出对称的统计特性,难以描述非对称性、有偏的输出噪声,使得在非对称偏斜噪声条件下算法的性能下降.基于此,研究一类广义双曲倾斜学生氏  $t$  (Generalized hyperbolic skew student's  $t$ , GHSkewt) 分布,并在非对称偏斜噪声条件下,提出一种线性系统鲁棒辨识算法.首先,对 GHSkewt 分布的重尾特性和偏斜特性进行详细阐述,数学上证明了标准学生氏  $t$  分布可看作是 GHSkewt 分布的一个特例;其次,引入隐含变量将 GHSkewt 分布进行数学分解,以方便算法的推导和实现;最后,在期望最大化 (Expectation-maximization, EM) 算法下,重构具有隐含变量系统的代价函数,通过迭代优化的方式,不断从被污染数据集中学习过程的动态特性和噪声分布,实现噪声参数和模型参数的联合估计.

**关键词** 鲁棒系统辨识,非对称偏斜噪声,广义双曲倾斜学生氏  $t$  分布,期望最大化算法

**引用格式** 刘鑫,陈强,王兰豪,代伟.非对称偏斜噪声条件下一种鲁棒概率系统辨识算法研究.自动化学报,2024,50(10): 2022–2035

DOI 10.16383/j.aas.c230624

CSTR 32138.14.j.aas.c230624

## Research on Robust Probabilistic System Identification Method With Asymmetric and Skewed Noise

LIU Xin<sup>1</sup> CHEN Qiang<sup>2</sup> WANG Lan-Hao<sup>3</sup> DAI Wei<sup>1,2</sup>

**Abstract** In the existing system identification algorithms, the commonly used Gaussian, student's  $t$  (St) and Laplace distributions all show symmetric statistical characteristics which makes them difficult to describe the asymmetric and skewed noise, therefore the performance of the corresponding algorithms may largely degrade with the skewed noise. To this end, this paper introduces the generalized hyperbolic skew student's  $t$  (GHSkewt) distribution and proposes a robust identification algorithm for linear systems with the asymmetric and skewed noise. Firstly, the thick-tailed and skewed characteristics of the GHSkewt distribution are introduced detailedly and it is also proved that the standard student's  $t$ -distribution can be regarded as a special case of the GHSkewt distribution; Secondly, the latent variables are introduced to mathematically decompose the GHSkewt distribution in order to facilitate the derivation and implementation of the algorithm; Finally, the system cost function with the latent variables is reconstructed under the expectation-maximization (EM) algorithm. The dynamic characteristics and noise distribution of the system are continuously learned from the contaminated data with iterative optimization, then the estimation of noise parameters and model parameters are realized.

**Key words** Robust system identification, asymmetric and skewed noise, generalized hyperbolic skew student's  $t$  (GHSkewt) distribution, expectation-maximization (EM) algorithm

**Citation** Liu Xin, Chen Qiang, Wang Lan-Hao, Dai Wei. Research on robust probabilistic system identification method with asymmetric and skewed noise. *Acta Automatica Sinica*, 2024, 50(10): 2022–2035

收稿日期 2023-10-09 录用日期 2024-03-10

Manuscript received October 9, 2023; accepted March 10, 2024

国家自然科学基金 (62103134, 62373361, 52304309), 国家重点研发计划 (2022YFB3304700), 中国博士后科学基金 (2023M743776) 资助

Supported by National Natural Science Foundation of China (62103134, 62373361, 52304309), National Key Research and Development Program of China (2022YFB3304700), and China Postdoctoral Science Foundation (2023M743776)

本文责任编辑 杨涛

Recommended by Associate Editor YANG Tao

1. 中国矿业大学人工智能研究院 徐州 221116 2. 中国矿业大学信息与控制工程学院 徐州 221116 3. 中国矿业大学炼焦煤资源绿色开发国家重点实验室 徐州 221116

1. Artificial Intelligence Research Institute, China University of

工业过程的复杂性使第一原理建模方法的使用受到限制,系统辨识作为一种过程数据驱动的建模方法,已在工业过程建模中受到了广泛关注<sup>[1-6]</sup>.与传统建模方法相比,系统辨识方法的优势在于无需探索系统内部的复杂机理,只需有效利用系统输入输出数据,便可简单、高效地实现系统等价模型的

Mining and Technology, Xuzhou 221116 2. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116 3. National Key Laboratory for Green Development of Coking Coal Resources, China University of Mining and Technology, Xuzhou 221116

构建<sup>[7-11]</sup>. 但在实际工业过程中, 数据采集过程未知的外部干扰、设备耦合、测量设备故障等因素常导致数据集中包含服从未知分布的噪声或异常值. 上述问题会严重降低辨识数据集的质量, 对系统辨识算法提出了更高要求. 在系统辨识领域, 由于数据集的质量不能得到稳定保证, 因此如何提高辨识算法的鲁棒性, 一直是研究的重点<sup>[12-15]</sup>. 阈值检测删除法是当输出数据超出一定阈值即判定为异常值, 仅保留非异常值数据用于参数估计. 显然, 数据点的删除会造成有用信息的缺失, 并且在复杂工业过程中选择适当阈值是困难的. 最早的鲁棒辨识算法是具有分段损失函数的 Huber 鲁棒回归. Huber 估计将最小绝对偏差损失函数应用于异常值, 将普通最小二乘方法应用于正常数据<sup>[16-18]</sup>.

鲁棒概率建模是降低异常值对算法性能影响的有效方法. 传统基于高斯分布推导的辨识算法为正常数据点和异常数据点分配等值权重是导致算法性能下降的本质原因<sup>[1, 17]</sup>. 高斯分布尾部较短, 对异常值缺乏鲁棒性, 无法降低异常值在算法参数估计中的权重. 针对这个问题, 文献 [19] 采用 2 个高斯分布混合对噪声建模, 以提高算法的鲁棒性; 利用方差较大的高斯分布对异常值建模, 利用方差较小的高斯分布拟合普通高斯白噪声, 在概率推导框架下得到分段线性自回归系统的鲁棒参数估计公式. 但是当异常值分布完全不规则时, 单个方差较大的高斯分布难以刻画其统计性质. 近年来, 研究者们注意到采用重尾分布建模噪声能够保障算法的鲁棒性. 常用重尾分布有学生氏  $t$  (Student's  $t$ ,  $St$ ) 分布、拉普拉斯 (Laplace) 分布及其多元形式. 这些重尾分布理论上能统计任意特征的异常值, 因为其在数学上能分解成无限子高斯分布的混合. 文献 [4] 针对线性自回归系统, 研究了基于学生氏  $t$  分布的鲁棒递推辨识算法. 进一步, 文献 [20] 在线性参数变化系统上, 基于学生氏  $t$  分布完成鲁棒概率建模, 提高了算法的鲁棒性; 文献 [21] 以状态空间模型为例, 采用学生氏  $t$  分布的重尾特性抑制异常值的影响, 提高了模型参数估计的精度; 文献 [22] 讨论了非线性状态空间的鲁棒辨识问题, 论述了采用学生氏  $t$  分布建模噪声的优越性, 并从数学角度对辨识算法的鲁棒性进行了清晰和明确的解释. 相比于学生氏  $t$  分布, 拉普拉斯分布的概率密度函数较为简单, 其有着更尖锐的峰值和更长尾部. 因此, 拉普拉斯分布驱动的鲁棒辨识算法在参数估计上有着更小的计算成本, 在算法收敛速度上有显著提高. 文献 [17] 为提高算法收敛速度, 采用拉普拉斯分布代替学生氏  $t$  分布统计系统的噪声特性, 提出一种线性系统

在线鲁棒辨识算法; 文献 [23] 针对工业过程数据存在复杂非高斯和多模态特征问题, 采用了多元拉普拉斯分布进行鲁棒建模.

上述鲁棒系统辨识算法都是基于对称重尾分布得到的, 利用分布的重尾特性保障算法的鲁棒性. 由于脉冲干扰、测量值异常和人为建模误差等因素, 非对称噪声在实际工业过程中普遍存在. 作为对称分布的推广, 非对称分布能够灵活地反映测量噪声中出现的偏态及重尾特性, 已在机电定位系统、超宽带测距定位、生物统计学和经济学等领域应用<sup>[24-31]</sup>. 如在编码器和驱动轴间的错误连接, 可导致机电定位系统中编码器的脉冲丢失, 从而引入位置偏差<sup>[24]</sup>; 在超宽带测距定位领域, 多径效应和非视距误差会导致非负的测量误差<sup>[31]</sup>. 因此, 在输出噪声服从非对称分布情况下, 现有基于对称分布建模的辨识算法估计误差会出现明显的偏斜, 导致算法性能下降. 为解决上述问题, 文献 [29] 研究了非对称测量噪声条件下线性状态空间系统辨识问题, 引入偏斜  $t$  分布描述数据呈现出的偏态和重尾特性, 以保证模型的精确度; 文献 [32] 研究了多模型的鲁棒线性参数变化模型, 采用非对称拉普拉斯分布描述实际工业过程非对称噪声. 上述工作进一步提高了算法的鲁棒性, 但所采用的分布较为固定, 难以统计实际过程中灵活多变的噪声分布特性.

广义双曲倾斜学生氏  $t$  (Generalized hyperbolic skew student's  $t$ , GHSkewt) 分布拥有着更多的参数调整空间, 其重尾和偏斜特性保障算法的鲁棒性, 并且能灵活调节其超参数, 以实现噪声分布特性的统计, 有着更广的适用范围<sup>[33-34]</sup>. 基于此, 本文考虑输出数据受到非对称偏斜噪声污染的问题, 有针对性地研究了 GHSkewt 分布的重尾特性和偏斜特性, 并基于此, 提出一种线性系统鲁棒辨识算法. 首先, 对 GHSkewt 分布进行详细介绍, 在数学上证明了标准学生氏  $t$  分布可以看作是 GHSkewt 分布的一个特例; 然后引入隐含变量, 将 GHSkewt 分布进行数学分解, 以方便算法的推导和实现; 最后, 在期望最大化 (Expectation-maximization, EM) 算法下, 重构具有隐含变量系统的代价函数, 通过迭代优化的方式, 不断从被污染数据集中, 学习过程的动态特性和噪声分布, 实现噪声参数和模型参数的联合估计. 本文主要贡献包括以下 3 点:

- 1) 考虑到输出数据中非对称偏斜测量噪声的污染问题, 系统地研究了 GHSkewt 分布的重尾特性和偏斜特性, 提出一种适用范围更广的鲁棒辨识算法;

- 2) 在数学上证明了标准学生氏  $t$  分布可以看作

是 GHSkewt 分布的一个特例, 并且引入隐含变量, 将 GHSkewt 分布分解, 使得算法的推导和实现更加高效;

3) 在 EM 算法框架下构建系统的概率模型, 算法能自适应地学习过程的动态特性和噪声分布, 以实现噪声参数和模型参数的联合估计.

## 1 辨识问题描述

### 1.1 系统模型

考虑下面的线性系统:

$$A(z^{-1})x_t = B(z^{-1})u_t + e_t \quad (1)$$

且

$$\begin{cases} A(z^{-1}) = 1 + a_1 z^{-1} + \dots + a_{n_a} z^{-n_a} \\ B(z^{-1}) = b_1 z^{-1} + \dots + b_{n_b} z^{-n_b} \end{cases} \quad (2)$$

式中,  $\{u_t, x_t, e_t\}_{t=1:N}$  分别表示系统的输入序列、输出序列和噪声序列,  $N$  为辨识数据的长度;  $A(z^{-1})$  和  $B(z^{-1})$  表示系统的多项式传递函数;  $z$  表示延迟算子, 即  $z^{-1}x_t = x_{t-1}$ ;  $n_a$  和  $n_b$  分别表示系统输出和输入传递函数的已知阶数.

结合式 (1) 和式 (2), 系统的模型可改写为:

$$x_t = \varphi_t^T \theta + e_t \quad (3)$$

式中

$$\varphi_t = [x_{t-1}, \dots, x_{t-n_a}, u_{t-1}, \dots, u_{t-n_b}]^T \quad (4)$$

待辨识的模型参数  $\theta$  可定义为:

$$\theta = [a_1, \dots, a_{n_a}, b_1, \dots, b_{n_b}]^T \quad (5)$$

### 1.2 广义双曲倾斜学生氏 t 分布

鲁棒辨识算法常采用重尾分布对输出噪声  $e_t$  建模, 以抑制异常值影响, 如学生氏 t 分布、拉普拉斯分布. 然而, 这些对称重尾分布无法满足偏斜噪声的统计需要, 一旦系统输出噪声  $e_t$  呈偏斜特性, 上述对称重尾分布驱动的鲁棒辨识算法会出现有偏参数估计, 导致模型精度下降. 基于此, 本文利用非对称且重尾的 GHSkewt 分布对噪声  $e_t$  建模, 进而提出一种适用范围更加广泛的鲁棒系统辨识算法.

为保障算法对偏斜噪声的有效性, 令噪声  $e_t$  服从 GHSkewt 分布<sup>[26]</sup>:

$$p(e_t | \beta, \Lambda, \nu) = \text{GHSkewt}(e_t | 0, \beta, \Lambda, \nu) =$$

$$2^{\frac{1-\nu}{2}} (\nu)^{\frac{\nu}{2}} |\beta|^{\frac{\nu+1}{2}} K_{\frac{1+\nu}{2}} \times \frac{\sqrt{\beta^2(\nu\Lambda + e_t^2)}}{\Lambda} \times \frac{\exp\left\{\frac{\beta e_t}{\Lambda}\right\}}{\left(\Gamma\left(\frac{\nu}{2}\right)\sqrt{\pi\Lambda}\left(\sqrt{\nu\Lambda + e_t^2}\right)\right)^{\frac{\nu+1}{2}}} \quad (6)$$

式中,  $\beta$ 、 $\Lambda$  和  $\nu$  分别表示 GHSkewt 分布的偏斜参数、尺度参数和自由度参数;  $p(\cdot)$  表示概率密度函数;  $\Gamma(\cdot)$  表示伽马函数;  $K_{(1+\nu)/2}(\cdot)$  表示第 2 类修正的贝塞尔函数, 其中  $(1+\nu)/2$  表示贝塞尔函数的阶次. 由式 (3) 可以看出, 系统输出  $x_t$  与噪声  $e_t$  有着相同的统计性质, 表示为:

$$p(x_t | \varphi_t^T \theta, \beta, \Lambda, \nu) = \text{GHSkewt}(x_t | \varphi_t^T \theta, \beta, \Lambda, \nu) \quad (7)$$

由于贝塞尔函数和伽马函数的存在, GHSkewt 分布复杂的表达式难以直接用于推导辨识算法. 为了便于后续分析和算法推导, 通过引入隐含变量, 可将 GHSkewt 分布分解成无数个次高斯分布的加权组合形式<sup>[26, 34]</sup>:

$$p(x_t | \varphi_t^T \theta, \beta, \Lambda, \nu) = \int N(x_t | \varphi_t^T \theta + \beta w_t, \Lambda w_t) p(w_t | \nu) dw_t \quad (8)$$

式中,  $\{w_t\}_{t=1:N}$  表示引入的未知隐含变量;  $N(x_t | \varphi_t^T \theta + \beta w_t, \Lambda w_t)$  表示 GHSkewt 分布的次高斯分布:

$$N(x_t | \varphi_t^T \theta + \beta w_t, \Lambda w_t) = \frac{1}{\sqrt{2\pi\Lambda w_t}} \exp\left\{-\frac{(x_t - \varphi_t^T \theta - \beta w_t)^2}{2\Lambda w_t}\right\} \quad (9)$$

概率密度函数  $p(w_t | \nu)$  可表示为:

$$p(w_t | \nu) = \text{IG}\left(w_t \mid \frac{\nu}{2}, \frac{\nu}{2}\right) = \frac{\left(\frac{\nu}{2}\right)^{\frac{\nu}{2}}}{\Gamma\left(\frac{\nu}{2}\right)} w_t^{-\frac{\nu}{2}-1} \exp\left\{-\frac{\nu}{2w_t}\right\} \quad (10)$$

式中,  $\text{IG}(w_t | \nu/2, \nu/2)$  表示概率密度函数为服从形状参数、尺度参数皆为  $\nu/2$  的逆伽马 (Inverse gamma, IG) 分布<sup>[34]</sup>.

由式 (8) 可以看出, GHSkewt 分布可看成具有变化均值和方差的无限次高斯分布的凸组合, 各次高斯分布的权重由服从逆伽马分布的  $p(w_t | \nu)$  决定. 下面将从数学上证明基于学生氏 t 分布的鲁棒辨识算法是基于 GHSkewt 分布的鲁棒辨识算法的一个特例, 由此说明基于 GHSkewt 分布的鲁棒辨识算法具有更广泛的应用范围.

假设式 (8) 中的偏斜参数  $\beta = 0$ , 这意味着输出测量噪声无偏斜性质, 此时系统输出  $x_t$  有下式成立:

$$p(x_t | \varphi_t^T \theta, 0, \Lambda, v) = \int N(x_t | \varphi_t^T \theta, \Lambda w_t) \text{IG}\left(w_t \left| \frac{v}{2}, \frac{v}{2} \right.\right) dw_t \quad (11)$$

令  $\lambda_t = 1/w_t$ , 在积分变量代换中积分限被改变, 为保持原有积分上下限, 式 (11) 可重写为:

$$p(x_t | \varphi_t^T \theta, 0, \Lambda, v) = \int N\left(x_t \left| \varphi_t^T \theta, \frac{\Lambda}{\lambda_t} \right.\right) \text{IG}\left(\frac{1}{\lambda_t} \left| \frac{v}{2}, \frac{v}{2} \right.\right) d\frac{1}{\lambda_t} \quad (12)$$

式中, 逆伽马分布  $\text{IG}\left(\frac{1}{\lambda_t} \left| \frac{v}{2}, \frac{v}{2} \right.\right)$  经数学变换, 可得:

$$\begin{aligned} \text{IG}\left(\frac{1}{\lambda_t} \left| \frac{v}{2}, \frac{v}{2} \right.\right) d\frac{1}{\lambda_t} &= \\ \frac{\left(\frac{v}{2}\right)^{\frac{v}{2}}}{\Gamma\left(\frac{v}{2}\right)} \lambda_t^{\frac{v}{2}+1} \exp\left\{-\frac{v\lambda_t}{2}\right\} d\frac{1}{\lambda_t} &= \\ -\frac{\left(\frac{v}{2}\right)^{\frac{v}{2}}}{\Gamma\left(\frac{v}{2}\right)} \lambda_t^{\frac{v}{2}-1} \exp\left\{-\frac{v\lambda_t}{2}\right\} d\lambda_t &= \\ -G\left(\lambda_t \left| \frac{v}{2}, \frac{v}{2} \right.\right) d\lambda_t & \quad (13) \end{aligned}$$

式中,  $G\left(\lambda_t \left| \frac{v}{2}, \frac{v}{2} \right.\right)$  表示概率密度函数为服从形状参数、尺度参数皆为  $v/2$  的伽马 (Gamma, G) 分布<sup>[34]</sup>.

结合式 (12) 和式 (13), 可得:

$$p(x_t | \varphi_t^T \theta, 0, \Lambda, v) = \int N\left(x_t \left| \varphi_t^T \theta, \frac{\Lambda}{\lambda_t} \right.\right) G\left(\lambda_t \left| \frac{v}{2}, \frac{v}{2} \right.\right) d\lambda_t \sim \text{St}(x_t | \varphi_t^T \theta, \Lambda, v) \quad (14)$$

式中,  $\text{St}(x_t | \varphi_t^T \theta, \Lambda, v)$  表示概率密度函数为服从均值为  $\varphi_t^T \theta$ 、尺度参数为  $\Lambda$  和自由度参数为  $v$  的学生氏  $t$  分布.

至此, 当 GHSkewt 分布的偏斜参数  $\beta = 0$  时, GHSkewt 分布退化为标准的学生氏  $t$  分布, 这意味着学生氏  $t$  分布是 GHSkewt 分布的一个特例. 从理论上可以得出, 针对输出异常值问题, GHSkewt 分布应具有与重尾学生氏  $t$  分布相同的鲁棒性<sup>[35]</sup>. 此外, 对比式 (8) 和式 (14) 可以看出, GHSkewt 分布的次高斯分布均值项在无噪声的系统输出  $\varphi_t^T \theta$  基础上引入可变化的偏斜量  $\beta w_t$ , 并且受权重影响, 这使得 GHSkewt 分布可以描述偏斜噪声, 从而消除对建模的负面影响.

为更直观地了解 GHSkewt 分布的统计特性, 在尺度参数  $\Lambda = 1$  时, 图 1(a) 和图 1(b) 分别绘制了式 (6) 当自由度参数  $v = 8$  和偏斜参数  $\beta = -4$  时, 不同偏斜参数  $\beta$  和自由度参数  $v$  的非对称 GH-Skewt 分布概率密度曲线, 并对比绘制具有对称特性的高斯分布  $N(\cdot)$ 、学生氏  $t$  分布  $\text{St}(\cdot)$  和拉普拉斯分布  $\text{Laplace}(\cdot)$  概率密度曲线. 由图 1 可以看出: 1) GHSkewt 分布具有偏斜和重尾特性, 当偏斜参数  $\beta$  和自由度参数  $v$  减小时, GHSkewt 分布的偏斜程度变大且尾部变长; 2) 当 GHSkewt 分布的偏斜参数  $\beta = 0$  时, GHSkewt 分布便退化为学生氏  $t$  分布, 这与上述理论证明吻合; 3) 当学生氏  $t$  分布的自由度参数趋于无穷大时, 学生氏  $t$  分布退化为高斯分布. 综上可以看出, GHSkewt 分布可看作是比学生氏  $t$  分布和高斯分布更具一般性的噪声分布, GHSkewt 分布除了具有描述高斯噪声和异常值的能力, 还具有描述偏斜噪声的能力.

在本文中, 可用数据集包括系统的输入输出数

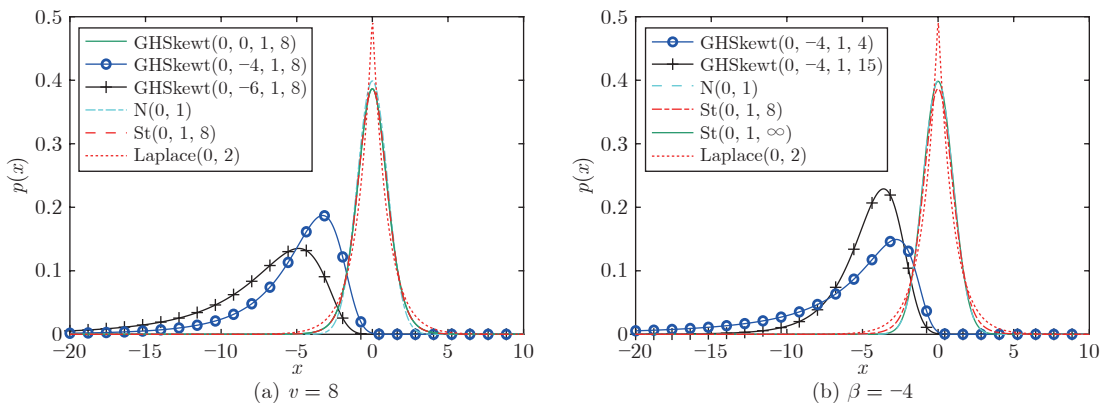


图 1 对称分布与参数值  $\beta$  和  $v$  不同的 GHSkewt 分布对比

Fig. 1 Comparison of the symmetric distribution and the GHSkewt distribution with different parameter values of  $\beta$  and  $v$

据  $U_{\text{obs}} = \{x_{1:N}, u_{1:N}\}$ , 因为引入了未知的隐含变量, 所以不可观测数据集可表示为  $U_{\text{mis}} = \{w_{1:N}\}$ ; 待辨的参数集包括未知模型参数和噪声参数, 可表示为  $P = \{\theta, \beta, \Lambda, v\}$ . 综上, 本文的主要任务可以概括为: 在系统输出被偏斜噪声污染的条件下, 基于可观测数据集  $U_{\text{obs}}$  估计未知参数集  $P$ .

## 2 基于 GHSkewt 分布的鲁棒辨识算法

### 2.1 EM 算法简介

EM 算法可以在具有未知变量的情况下完成参数的极大似然估计, 故 EM 算法在系统辨识领域得到了非常广泛的应用, 尤其是针对数据丢失和隐含变量问题<sup>[9]</sup>. EM 算法实际上是一种迭代优化算法, 通过重复执行期望步骤 (E-step) 和最大化步骤 (M-step), 以实现模型参数的估计. 在 E-step 中, 首先构造系统对数似然函数:

$$L = \ln p(U_{\text{obs}}, U_{\text{mis}}|P) \quad (15)$$

由于存在不可观测数据集  $U_{\text{mis}}$ , 直接优化上述函数无法实现. 在 E-step 中, 通过计算上述对数似然函数相对于隐藏变量的期望, 来构造如下目标代价函数:

$$Q(P|P^m) = E_{U_{\text{mis}}|U_{\text{obs}}, P^m} [\ln p(U_{\text{obs}}, U_{\text{mis}}|P)] = \int p(U_{\text{mis}}|U_{\text{obs}}, P^m) \ln p(U_{\text{obs}}, U_{\text{mis}}|P) dU_{\text{mis}} \quad (16)$$

式中,  $P^m$  表示在第  $m$  次迭代中得到的参数估计值. 接着在 M-step 中, 通过最大化目标代价函数得到新的未知参数估计值:

$$P^{m+1} = \arg \max_P Q(P|P^m) \quad (17)$$

重复执行 E-step 和 M-step, 直到待辨识参数收敛得到最优估计. 算法 1 为 EM 算法的具体步骤.

#### 算法 1. EM 算法

**初始化.** 待辨识参数集  $P^1$  且设置  $m = 1$ ;

- 1) E-step: 根据式 (16), 计算目标代价函数;
- 2) M-step: 根据式 (17), 最大化目标代价函数, 得到新的参数估计值  $P^{m+1}$ ;
- 3) 令  $m \leftarrow m + 1$  并重复步骤 2) 和步骤 3), 直到参数收敛.

### 2.2 基于 GHSkewt 分布的鲁棒辨识算法数学推导

#### 1) E-step

为构造目标代价函数, 首先根据贝叶斯定理, 将系统的对数似然函数分解如下:

$$L = \ln p(U_{\text{obs}}, U_{\text{mis}}|P) = \ln p(x_{1:N}, u_{1:N}, w_{1:N}|P) = \ln p(x_{1:N}|u_{1:N}, w_{1:N}, P) + \ln p(w_{1:N}|u_{1:N}, P) + C_1 \quad (18)$$

式中,  $C_1 = \ln p(u_{1:N}|P)$  表示与参数无关的常数项. 基于式 (3) 和式 (4), 概率密度函数  $\ln p(x_{1:N}|u_{1:N}, w_{1:N}, P)$  可进一步简化为:

$$\ln p(x_{1:N}|u_{1:N}, w_{1:N}, P) = \ln p(x_N, x_{1:N-1}|u_{1:N}, w_{1:N}, P) = \ln p(x_N|x_{1:N-1}, u_{1:N}, w_{1:N}, P) + \ln p(x_{1:N-1}|u_{1:N}, w_{1:N}, P) = \sum_{t=1}^N \ln p(x_t|\varphi_t, w_t, P) \quad (19)$$

由于未知隐含变量  $w_t$  与系统输入输出无关, 因此  $\ln p(w_{1:N}|u_{1:N}, P)$  可进一步简化为:

$$\ln p(w_{1:N}|u_{1:N}, P) = \sum_{t=1}^N \ln p(w_t|P) \quad (20)$$

结合式 (9) 和式 (10), 系统对数似然函数可改写为:

$$L = \ln p(x_{1:N}, u_{1:N}, w_{1:N}|P) = \sum_{t=1}^N \ln p(x_t|\varphi_t, w_t, P) + \sum_{t=1}^N \ln p(w_t|P) + C_1 = \sum_{t=1}^N \ln N(x_t|\varphi_t^T \theta + \beta w_t, \Lambda w_t) + \sum_{t=1}^N \ln \text{IG}\left(w_t \left| \frac{v}{2}, \frac{v}{2} \right.\right) + C_1 \quad (21)$$

再基于式 (16), 计算系统对数似然函数  $L$  的条件期望, 得到目标代价函数为:

$$Q(P|P^m) = \sum_{t=1}^N \int_{w_t} [p(w_t|U_{\text{obs}}, P^m) \times \ln N(x_t|\varphi_t^T \theta + \beta w_t, \Lambda w_t)] dw_t + \sum_{t=1}^N \int_{w_t} p(w_t|U_{\text{obs}}, P^m) \ln \text{IG}\left(w_t \left| \frac{v}{2}, \frac{v}{2} \right.\right) dw_t + C_1 \quad (22)$$

根据式 (9) 和式 (10), 将式 (22) 中高斯分布  $N(\cdot)$  和逆伽马分布  $\text{IG}(\cdot)$  的概率密度函数展开, 目标代价函数可写为:

$$\begin{aligned}
Q(P|P^m) = & -\frac{N}{2} \ln \Lambda - \frac{1}{2\Lambda} \sum_{t=1}^N (x_t - \varphi_t^T \boldsymbol{\theta})^2 \langle w_t^{-1} \rangle + \\
& \frac{\beta}{\Lambda} \sum_{t=1}^N (x_t - \varphi_t^T \boldsymbol{\theta}) - \frac{\beta^2}{2\Lambda} \sum_{t=1}^N \langle w_t \rangle + \\
& \frac{N\nu}{2} \ln \frac{\nu}{2} - N \ln \Gamma \left( \frac{\nu}{2} \right) - \frac{\nu}{2} \sum_{t=1}^N \langle w_t^{-1} \rangle - \\
& \left( \frac{\nu}{2} + 1 \right) \sum_{t=1}^N \langle \ln w_t \rangle + C_2 \quad (23)
\end{aligned}$$

式中,  $C_2 = C_1 - (N/2) \ln 2\pi - (1/2) \sum_{t=1}^N \langle \ln w_t \rangle$  表示与参数无关项;  $\langle w_t \rangle$ 、 $\langle w_t^{-1} \rangle$  和  $\langle \ln w_t \rangle$  分别表示有关隐含变量  $w_t$  的 3 种期望项:

$$\langle w_t \rangle = \int p(w_t | x_{1:N}, u_{1:N}, P^m) w_t dw_t \quad (24)$$

$$\langle w_t^{-1} \rangle = \int p(w_t | x_{1:N}, u_{1:N}, P^m) w_t^{-1} dw_t \quad (25)$$

$$\langle \ln w_t \rangle = \int p(w_t | x_{1:N}, u_{1:N}, P^m) \ln w_t dw_t \quad (26)$$

式中,  $p(w_t | x_{1:N}, u_{1:N}, P^m)$  表示隐含变量  $w_t$  在第  $m$  次参数估计  $P^m$  和可观测数据集下的后验概率密度. 根据贝叶斯条件概率公式, 可得:

$$\begin{aligned}
p(w_t | x_{1:N}, u_{1:N}, P^m) = & \frac{p(w_t, x_t | y_{1:t-1}, u_{1:t-1}, P^m)}{p(x_t | x_{1:t-1}, u_{1:t-1}, P^m)} = \\
& \frac{p(x_t | \varphi_t^T, w_t, P^m) p(w_t | P^m)}{p(x_t | \varphi_t, P^m)} = \\
& \frac{N(x_t | \varphi_t^T \boldsymbol{\theta}^m + \beta^m w_t, \Lambda^m w_t) \text{IG}(w_t | \frac{\nu^m}{2}, \frac{\nu^m}{2})}{\text{GHSkewt}(x_t | \varphi_t^T \boldsymbol{\theta}^m, \beta^m, \Lambda^m, \nu^m)} = \\
& \text{GIG}(w_t | \lambda, \delta, \gamma) \quad (27)
\end{aligned}$$

式中,  $\text{GIG}(w_t | \lambda, \delta, \gamma)$  表示  $p(w_t | x_{1:N}, u_{1:N}, P^m)$  是服从参数为  $\lambda$ 、 $\delta$  和  $\gamma$  的广义逆高斯 (Generalized inverse Gaussian, GIG) 分布<sup>[34]</sup>. 3 个参数具体计算公式为:

$$\lambda = -\frac{\nu^m + 1}{2} \quad (28)$$

$$\delta = \sqrt{\nu^m + \frac{(x_t - \varphi_t^T \boldsymbol{\theta}^m)^2}{\Lambda^m}} \quad (29)$$

$$\gamma = \sqrt{\frac{(\beta^m)^2}{\Lambda^m}} \quad (30)$$

根据文献 [26], 由式 (24) ~ (26), 可得期望项的计算表达式为:

$$\langle w_t \rangle = \left( \frac{\delta}{\gamma} \right) \frac{K_{\lambda+1}(\delta\gamma)}{K_{\lambda}(\delta\gamma)} \quad (31)$$

$$\langle w_t^{-1} \rangle = \left( \frac{\delta}{\gamma} \right)^{-1} \frac{K_{\lambda-1}(\delta\gamma)}{K_{\lambda}(\delta\gamma)} \quad (32)$$

$$\langle \ln w_t \rangle = \ln \left( \frac{\delta}{\gamma} \right) + \frac{1}{K_{\lambda}(\delta\gamma)} \frac{\partial}{\partial \lambda} K_{\lambda}(\delta\gamma) \quad (33)$$

至此, 在 E-step 中, 目标代价函数构造完成. 为了表述清晰, 可将目标代价函数  $Q(P|P^m)$  拆分为以下形式:

$$Q(P|P^m) = Q_1(\boldsymbol{\theta}, \beta, \Lambda) + Q_2(\nu) + C_2 \quad (34)$$

式中

$$\begin{aligned}
Q_1(\boldsymbol{\theta}, \beta, \Lambda) = & -\frac{N}{2} \ln \Lambda - \frac{1}{2\Lambda} \sum_{t=1}^N (x_t - \varphi_t^T \boldsymbol{\theta})^2 \langle w_t^{-1} \rangle + \\
& \frac{\beta}{\Lambda} \sum_{t=1}^N (x_t - \varphi_t^T \boldsymbol{\theta}) - \frac{\beta^2}{2\Lambda} \sum_{t=1}^N \langle w_t \rangle \quad (35)
\end{aligned}$$

$$\begin{aligned}
Q_2(\nu) = & \frac{N\nu}{2} \ln \frac{\nu}{2} - N \ln \Gamma \left( \frac{\nu}{2} \right) - \frac{\nu}{2} \sum_{t=1}^N \langle w_t^{-1} \rangle - \\
& \left( \frac{\nu}{2} + 1 \right) \sum_{t=1}^N \langle \ln w_t \rangle + C_2 \quad (36)
\end{aligned}$$

## 2) M-step

在 M-step 中, 基于式 (17), 通过最大化式 (34) ~ (36) 中的目标代价函数  $Q(P|P^m)$ , 以得到未知参数估计.

首先, 通过优化  $Q_1(\boldsymbol{\theta}, \beta, \Lambda)$  以求解模型参数  $\boldsymbol{\theta}$ 、噪声偏斜参数  $\beta$  和尺度参数  $\Lambda$  的迭代估计公式. 需要注意的是,  $Q_1(\boldsymbol{\theta}, \beta, \Lambda)$  是一个复合函数, 这里需要采用 3 步优化法, 即当求解  $\boldsymbol{\theta}$  的估计公式时, 令  $\beta = \beta^m$  和  $\Lambda = \Lambda^m$  为定值; 当求解  $\beta$  的估计公式时, 令  $\boldsymbol{\theta} = \boldsymbol{\theta}^m$  和  $\Lambda = \Lambda^m$  为定值; 当求解  $\Lambda$  的估计公式时, 令  $\boldsymbol{\theta} = \boldsymbol{\theta}^m$  和  $\beta = \beta^m$  为定值. 计算  $Q_1(\boldsymbol{\theta}, \beta, \Lambda)$  对  $\boldsymbol{\theta}$  的导数并令其等于零, 可得:

$$\begin{aligned}
\frac{\partial Q_1(\boldsymbol{\theta}, \beta, \Lambda)}{\partial \boldsymbol{\theta}} = 0 \quad \Rightarrow \quad \boldsymbol{\theta}^{m+1} = & \frac{\sum_{t=1}^N (\varphi_t x_t \langle w_t^{-1} \rangle - \beta^m \varphi_t)}{\sum_{t=1}^N (\varphi_t \varphi_t^T \langle w_t^{-1} \rangle)} \quad (37)
\end{aligned}$$

然后, 求解  $Q_1(\boldsymbol{\theta}, \beta, \Lambda)$  对  $\beta$  的导数并令其等于零, 可得:

$$\frac{\partial Q_1(\boldsymbol{\theta}, \beta, \Lambda)}{\partial \beta} = 0 \Rightarrow \beta^{m+1} = \frac{\sum_{t=1}^N (x_t - \boldsymbol{\varphi}_t^T \boldsymbol{\theta}^m)}{\sum_{t=1}^N \langle w_t \rangle} \quad (38)$$

最后, 求解  $Q_1(\boldsymbol{\theta}, \beta, \Lambda)$  对  $\Lambda$  的导数并令其等于零, 可得:

$$\frac{\partial Q_1(\boldsymbol{\theta}, \beta, \Lambda)}{\partial \Lambda} = 0 \Rightarrow \Lambda^{m+1} = \frac{1}{N} \left[ \sum_{t=1}^N (x_t - \boldsymbol{\varphi}_t^T \boldsymbol{\theta}^m)^2 \langle w_t^{-1} \rangle + (\beta^m)^2 \sum_{t=1}^N \langle w_t \rangle - 2\beta^m \sum_{t=1}^N (x_t - \boldsymbol{\varphi}_t^T \boldsymbol{\theta}^m) \right] \quad (39)$$

为了求解自由度参数  $v$  的估计公式, 计算函数  $Q_2(v)$  对  $v$  的导数并令其等于零, 可得:

$$\frac{\partial Q_2(v)}{\partial v} = 0 \Rightarrow \frac{N}{2} \ln \frac{v}{2} + \frac{N}{2} - \frac{N}{2} \psi \left( \frac{v}{2} \right) - \frac{1}{2} \sum_{t=1}^N \langle w_t^{-1} \rangle - \frac{1}{2} \sum_{t=1}^N \langle \ln w_t \rangle = 0 \quad (40)$$

式中,  $\psi(x) = \frac{d \ln \Gamma(x)}{dx}$  表示对数伽马函数的 1 阶导数. 求解式 (40), 可以得到自由度参数  $v$  的估计结果, 显然式 (40) 的解析解并不存在, 因此需要通过数值优化的方法优化式 (40), 以得到参数  $v$  的估计值. 通常使用 Matlab 中的 `fsolve` 函数求解式 (40), 得到自由度参数  $v$  的优化估计值.

至此, 在偏斜噪声条件下基于 GHSkewt 分布的鲁棒辨识算法推导完成. 算法 2 为本文提出的基于 GHSkewt 分布的鲁棒辨识算法.

### 算法 2. 基于 GHSkewt 分布的鲁棒辨识算法

**输入.** 可观测数据集  $U_{\text{obs}} = \{x_{1:N}, u_{1:N}\}$ .

**输出.** 最终迭代结果为较好估计参数  $P^* = \{\boldsymbol{\theta}^*, \beta^*, \Lambda^*, v^*\}$ .

1) 初始化  $P^m = P^1 = \{\boldsymbol{\theta}^1, \beta^1, \Lambda^1, v^1\}$  且  $m = 1$ .

E-step:

2) 根据式 (27), 计算隐含变量  $w_t$  的后验概率密度函数

$$p(w_t | x_{1:N}, u_{1:N}, P^m);$$

3) 根据式 (31) ~ (33), 分别计算必要的期望项  $\langle w_t \rangle$ 、 $\langle w_t^{-1} \rangle$  和  $\langle \ln w_t \rangle$ ;

4) 根据式 (34) ~ (36), 构造目标代价函数  $Q(P|P^m)$ .

M-step:

5) 根据式 (37) ~ (39), 分别计算参数  $\boldsymbol{\theta}$ 、 $\beta$  和  $\Lambda$ ;

6) 通过优化式 (40), 更新自由度参数  $v$ ;

7) 如果  $\frac{\|\boldsymbol{\theta}^{m+1} - \boldsymbol{\theta}^m\|_2}{\|\boldsymbol{\theta}^m\|_2} \leq 10^{-3}$ , 算法运行终止; 否则, 令  $m \leftarrow m + 1$  并重复步骤 2) ~ 6), 直到满足终止条件.

针对算法 2 参数初值选取, 本文在初值随机选取的多次预运行实验下确定参数初值  $P^1$ . 每次预运行实验的停止准则为:

$$\frac{\ln p(U|P^m) - \ln p(U|P^{m-1})}{\ln p(U|P^m) - \ln p(U|P^0)} \leq 10^{-2} \quad (41)$$

式中,  $U = \{U_{\text{obs}}, U_{\text{mis}}\}$  表示完整数据集,  $m$  表示迭代次数,  $P^0$  表示预运行实验随机初值. 每次预运行实验结束后, 令  $\tilde{P}^* = P^m$ , 经过  $h$  次预运行实验后, 算法 2 的初值选取为:

$$P^1 = \arg \max_{\tilde{P}_i^*} \ln p(U|\tilde{P}_i^*), \quad i = 1, 2, \dots, h \quad (42)$$

## 3 算法验证

本节将系统地验证算法 2 的有效性. 因为本文的算法 2 是基于 GHSkewt 分布提出的, 为方便描述, 将算法 2 简称为 GHSkewt-Iden. 文献 [4, 10] 基于学生氏  $t$  分布和拉普拉斯分布提出 2 种鲁棒线性系统辨识算法, 分别简称为 St-Iden 和 Laplace-Iden. 为了更好地体现本文 GHSkewt-Iden 算法的有效性和先进性, 将 St-Iden 算法和 Laplace-Iden 算法设置为对比算法, 在公平的实验条件下设计对比实验. 为了更清楚地评估算法的性能, 采用相对参数估计误差 (Relative parameter estimation error, RPEE)、均方根误差 (Root mean square error, RMSE) 和决定系数 (Coefficient of determination,  $R^2$ ) 作为评价指标, 分别定义如下:

$$\text{RPEE} = \frac{\|\boldsymbol{\theta}^* - \boldsymbol{\theta}\|_2}{\|\boldsymbol{\theta}\|_2} \times 100\% \quad (43)$$

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{t=1}^N (x_t - \hat{x}_t)^2} \quad (44)$$

$$R^2 = 1 - \frac{\sum_{t=1}^N (x_t - \hat{x}_t)^2}{\sum_{t=1}^N (x_t - \bar{x})^2} \quad (45)$$

式中,  $\boldsymbol{\theta}^*$  和  $\boldsymbol{\theta}$  分别表示参数的估计值和真实值,  $x_t$  和  $\hat{x}_t$  分别表示第  $t$  个采样时刻的真实输出和估计输出;  $\bar{x}$  表示所有真实输出数据的平均值, 定义如下:

$$\bar{x} = \frac{1}{N} \sum_{t=1}^N x_t \quad (46)$$

### 3.1 数值例子

在线性自回归系统式 (1) 中:

$$A(z^{-1}) = 1 - 0.6z^{-1} + 0.8z^{-2} \quad (47)$$

$$B(z^{-1}) = 0.5z^{-1} + 0.4z^{-2} \quad (48)$$

模型参数真值:

$$\theta = [a_1, a_2, b_1, b_2]^T = [-0.6, 0.8, 0.5, 0.4]^T \quad (49)$$

为收集辨识数据集, 将均匀分布在  $-1 \sim 1$  的随机信号选取为输入信号用于激励系统, 数据长度设置  $N = 500$ . 为了验证本文算法的有效性, 进行 2 种测试: 实验 1 测试本文 GHSkewt-Iden 算法处理输出异常值的鲁棒性, 实验 2 测试本文 GHSkewt-Iden 算法处理偏斜噪声的有效性.

#### 实验 1. 输出异常值的鲁棒性测试

为了验证本文 GHSkewt-Iden 算法处理输出异常值的鲁棒性, 将在  $[-1.5, 1.5]$  均匀分布的随机数当作异常值加入真实输出数据中. 当异常值比例为 15% 时, 所得到的真实输出和含 15% 异常值比例输出如图 2 所示.

分别在输出异常值比例为 5%、10%、15%、20% 条件下, 针对 GHSkewt-Iden、St-Iden 和 Laplace-Iden 算法执行蒙特卡洛对比实验. 在每次蒙特卡洛实验中, 各算法独立运行 100 次, 这意味着蒙特卡洛仿真实验将得到 100 个模型的参数估计值. 在异常值比例为 15% 条件下, 4 个算法参数估

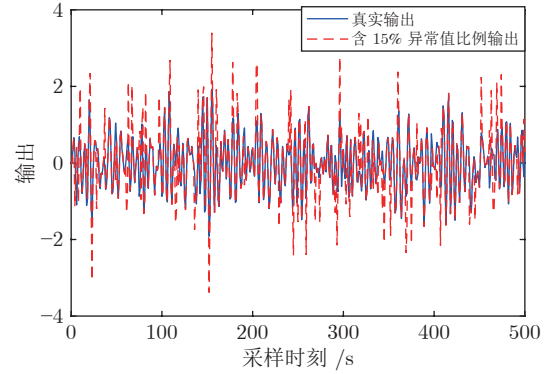


图 2 真实输出和含 15% 异常值比例输出的对比

Fig.2 Comparison of the real output and output with 15% of outlier points

计的均值和标准差如表 1 所示. 由表 1 可以看出, GHSkewt-Iden、St-Iden 和 Laplace-Iden 算法得到的模型参数均值都在真实值附近, 这表明本文 GHSkewt-Iden 算法对异常值同样具有鲁棒性.

此外, 为了进一步测试本文 GHSkewt-Iden 算法的性能. 在每次蒙特卡洛实验中, 根据得到的 100 个模型参数估计值, 各计算 100 次, 得到 RPEE、RMSE 和  $R^2$  指标, 并分别取其平均值, 统计结果见表 2. 由表 2 可以看出, 在参数估计的精确性和输出预测的准确性度量方面, 本文的 GHSkewt-Iden 算法与 St-Iden、Laplace-Iden 算法具有可比性, 这进一步验证了本文 GHSkewt-Iden 算法对异常值的有效性.

综上, 可以得出以下结论:

1) 在不同异常值比例条件下, 本文 GHSkewt-

表 1 异常值比例为 15% 时, 蒙特卡洛仿真实验参数估计结果的均值和标准差

Table 1 The mean and standard deviation of the Monte Carlo parameter estimation results with 15% of outlier points

算法	参数 $a_1$		参数 $a_2$		参数 $b_1$		参数 $b_2$	
	均值	标准差	均值	标准差	均值	标准差	均值	标准差
Laplace-Iden	-0.5751	$3.6 \times 10^{-4}$	0.7739	$1.9 \times 10^{-4}$	0.5083	$2.8 \times 10^{-4}$	0.4085	$2.7 \times 10^{-4}$
St-Iden	-0.5878	$3.2 \times 10^{-8}$	0.7899	$1.1 \times 10^{-8}$	0.5043	$6.5 \times 10^{-8}$	0.4028	$2.0 \times 10^{-8}$
GHSkewt-Iden	-0.5876	$6.5 \times 10^{-7}$	0.7899	$2.5 \times 10^{-7}$	0.5042	$1.1 \times 10^{-7}$	0.4030	$4.0 \times 10^{-7}$
真实值	-0.6000	—	0.8000	—	0.5000	—	0.4000	—

表 2 不同异常值比例下蒙特卡洛仿真实验的平均 RPEE、RMSE 和  $R^2$

Table 2 The averaged RPEE, RMSE and  $R^2$  for the Monte Carlo with different ratios of outlier points

异常值比例	Laplace-Iden			St-Iden			GHSkewt-Iden		
	RPEE (%)	RMSE	$R^2$	RPEE (%)	RMSE	$R^2$	RPEE (%)	RMSE	$R^2$
5%	0.71	0.0144	0.9995	0.61	0.0104	0.9998	0.39	0.0072	0.9999
10%	2.44	0.0475	0.9949	1.20	0.0186	0.9992	1.13	0.0192	0.9992
15%	3.20	0.0564	0.9928	1.40	0.0267	0.9984	1.42	0.0271	0.9983
20%	5.33	0.0880	0.9826	2.70	0.0499	0.9944	2.74	0.0503	0.9943



Iden 算法所得到的蒙特卡洛均值和标准差与 St-Iden、Laplace-Iden 算法具有可比性, 且基于 3 种辨识算法得到的蒙特卡洛均值稳定在真值附近, 可以给出精度较高的参数估计和模型预测结果. 这验证了本文 GHSkewt-Iden 算法对输出异常值具有鲁棒性.

2) 随着异常值比例增大和数据质量变差, 本文 GHSkewt-Iden 算法的参数估计精度和模型预测准确度有所下降. 这说明本文的 GHSkewt-Iden 算法本质上是一种基于过程数据的辨识算法, 其算法性能受数据质量影响.

**实验 2. 偏斜噪声的有效性测试**

在系统输出受到偏斜噪声污染条件下, 对本文 GHSkewt-Iden 算法的有效性进行测试. 首先, 生成尺度、偏斜和自由度参数分别为 0.000 1、-0.15 和 20 的服从 GHSkewt 分布的偏斜噪声, 记为:

$$\mathbf{G}_0 = [0.000\ 1, -0.15, 20] \quad (50)$$

偏斜噪声  $\mathbf{G}_0$  的概率密度如图 3 所示. 为了测试各算法对偏斜噪声的有效性, 将生成的偏斜噪声加入系统真实输出数据中, 分别执行 GHSkewt-Iden、St-Iden 和 Laplace-Iden 算法, 得到 3 种辨识算法在偏斜噪声  $\mathbf{G}_0$  下的参数估计轨迹如图 4 所示.

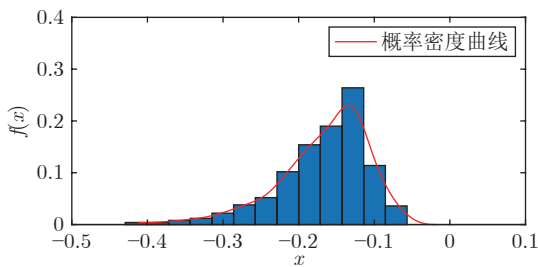


图 3 偏斜噪声  $\mathbf{G}_0$  的概率密度曲线

Fig.3 The probability density curve of skewed noise  $\mathbf{G}_0$

根据 3 种算法的参数估计结果, 分别计算 RP-EE、RMSE 和  $R^2$  指标, 对比结果如图 5 所示. 由图 4 和图 5 可以看出, 在输出数据中存在偏斜噪声条件下, 本文 GHSkewt-Iden 算法的参数估计精度明显高于 St-Iden 和 Laplace-Iden 算法. 图 6 给出了在本文 GHSkewt-Iden 算法执行过程中代价函数的收敛曲线, 验证了本文 GHSkewt-Iden 算法的收敛性.

在不同偏斜噪声条件下, 分别执行 GHSkewt-Iden、St-Iden 和 Laplace-Iden 算法进行对比, 以验证 GHSkewt-Iden 算法对偏斜噪声的有效性并非偶然. 首先选取 3 类服从 GHSkewt 分布的偏斜噪声参数为:

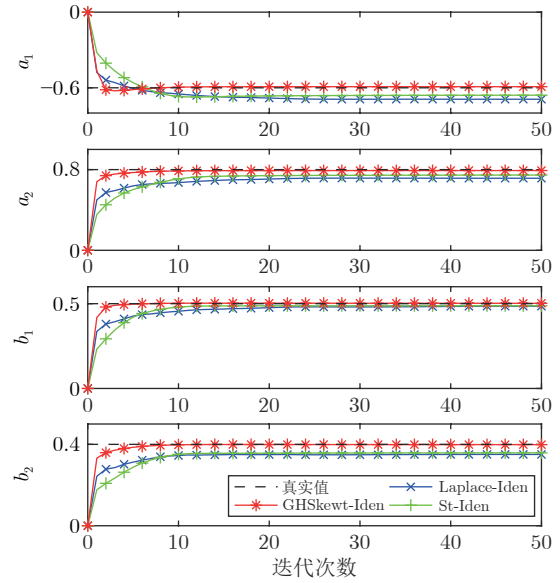


图 4 偏斜噪声  $\mathbf{G}_0$  下参数估计轨迹

Fig.4 Trajectories of the parameter estimates with skewed noise  $\mathbf{G}_0$

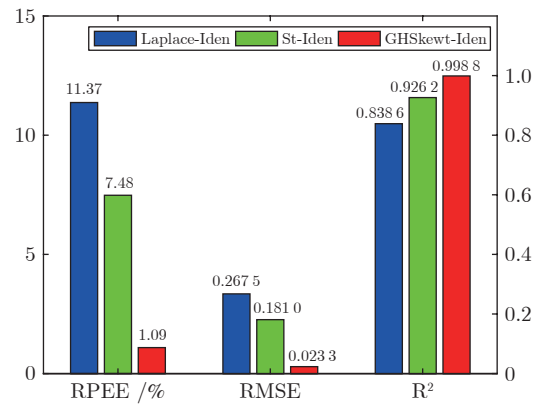


图 5 偏斜噪声  $\mathbf{G}_0$  下 3 种算法实验结果的对比

Fig.5 Comparison of experimental results of three algorithms under skewed noise  $\mathbf{G}_0$

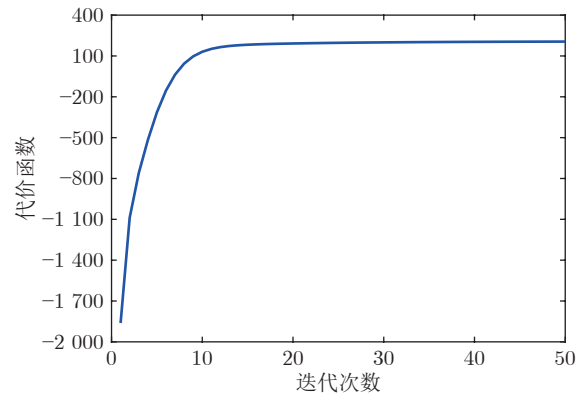


图 6 本文算法的代价函数轨迹

Fig.6 Trajectory of the cost function for the proposed algorithm

$$\begin{cases} \mathbf{G}_1 = [0.001, -0.15, 20] \\ \mathbf{G}_2 = [0.001, -0.20, 15] \\ \mathbf{G}_3 = [0.001, -0.15, 10] \end{cases} \quad (51)$$

首先, 将式 (51) 的 3 组参数生成的 3 种偏斜噪声加入真实输出数据中; 然后, 分别对 GHSkewt-Iden、St-Iden 和 Laplace-Iden 算法执行蒙特卡洛仿真实验; 最后, 计算各蒙特卡洛仿真参数估计的平均 RPEE、RMSE 和  $R^2$  指标, 结果如表 3 所示. 再根据 RPEE 指标的计算结果, 绘制了 3 种算法在不同偏斜噪声下蒙特卡洛仿真实验的 RPEE 曲面图, 如图 7 所示. 由表 3 可以看出, 在不同偏斜噪声条件下, 本文 GHSkewt-Iden 算法的 3 种性能指标明显优于其他 2 种算法, 验证了本文 GHSkewt-Iden 算法的先进性.

为了测试本文 GHSkewt-Iden 算法对初始参数的敏感性, 设计了不同参数初始值条件下的蒙特卡洛仿真实验. 首先, 在输出数据中加入偏斜噪声, 该偏斜噪声服从参数为  $[0.001, -0.15, 20]$  的 GH-Skewt 分布. 借鉴文献 [9], 在蒙特卡洛仿真实验中, 参数的初始值都从以真实参数为中心的对称区间中均匀选取:

$$\begin{aligned} \theta^0 &= [a_1^0, a_2^0, b_1^0, b_2^0]^T = \\ &[R(-1.2, 0), R(0.2, 1.4), R(0, 1), R(0, 0.8)]^T \end{aligned} \quad (52)$$

式中,  $R$  表示均匀分布.

图 8 为 100 次蒙特卡洛仿真参数估计的收敛轨迹, 其中红色虚线表示真实值, 蓝色区域为 100 次蒙特卡洛仿真参数估计运行得到的参数估计收敛曲线. 100 次蒙特卡洛仿真参数估计的均值和标准差如表 4 所示. 由图 8 可以看出, 在不同初始值条件下, 基于本文 GHSkewt-Iden 算法得到的模型参数均能收敛到真实值附近; 由表 4 可以看出, 参数的蒙特卡洛均值非常接近真实值且标准差均趋近于零, 这表明算法在准确估计模型参数的同时, 还具有很强的稳定性.

综上, 根据上述辨识结果可以看出:

1) 由图 4 可以看出, 在偏斜噪声条件下, 本文

GHSkewt-Iden 算法提供更精确的参数估计且收敛所需的迭代次数较少, 而 St-Iden 和 Laplace-Iden 算法给出的参数估计精度较低, 甚至出现了有偏的参数估计结果; 由图 5 可以看出, 在偏斜噪声条件下, 基于本文 GHSkewt-Iden 算法得到的模型预测精度明显高于 St-Iden 和 Laplace-Iden 算法. 这说明针对偏斜噪声问题, 本文 GHSkewt-Iden 算法比 St-Iden 和 Laplace-Iden 算法更具有鲁棒性.

2) 由图 7 可以看出, 在不同偏斜噪声条件下的蒙特卡洛仿真实验中, 基于本文 GHSkewt-Iden 算法得到的参数估计结果更接近真实值; 由表 3 可以看出, 基于本文 GHSkewt-Iden 算法得到的模型输出预测准确性更高. 这说明本文提出的 GHSkewt-Iden 算法不仅能有效地处理偏斜噪声问题, 给出精确度较高的模型参数估计, 同时算法具有稳定性, 在不同数据条件下均能给出有效的估计结果.

3) 由图 8 和表 4 可以看出, 基于本文 GHSkewt-Iden 算法在不同初始值条件下, 均能精准地收敛到真实值附近, 同时计算得到的标准差较小; 图 6 为在偏斜噪声条件下的代价函数轨迹.

上述辨识结果说明了针对偏斜噪声问题, 本文 GHSkewt-Iden 算法具有稳定性和收敛性.

### 3.2 质量弹簧阻尼系统

在质量弹簧阻尼系统上, 验证本文 GHSkewt-Iden 算法的有效性. 质量弹簧阻尼系统是工程领域常见的机械装置, 其结构如图 9 所示. 在该系统中, 质量块与 1 个弹簧和 1 个定阻尼器连接. 当质量块的质量很小时, 该系统的动力学方程表达式为:

$$\frac{d}{dt} \left( m \frac{d}{dt} S(t) \right) = F(t) - k_s S(t) - c_d \frac{d}{dt} S(t) \quad (53)$$

式中,  $S(t)$  表示质量块的垂直位置,  $k_s$  表示弹簧的刚度,  $c_d$  表示定阻尼器的阻尼系数,  $F(t)$  表示施加在质量块  $m$  上的外力.

在上述质量弹簧阻尼系统中, 弹簧的刚度设置  $k_s = 0.85 \text{ N/m}$ , 质量块的重量设置  $m = 0.01 \text{ kg}$ , 阻尼系数设置  $c_d = 0.5 \text{ N} \cdot \text{m}^2$ . 由于外力  $F(t)$  变化会影响质量块在垂直方向上的位置  $S(t)$ , 因此选取

表 3 不同偏斜噪声下蒙特卡洛仿真的平均 RPEE、RMSE 和  $R^2$

Table 3 The averaged RPEE, RMSE and  $R^2$  for the Monte Carlo with different levels of skewed noise

偏斜噪声	Laplace-Iden			St-Iden			GHSkewt-Iden		
	RPEE (%)	RMSE	$R^2$	RPEE (%)	RMSE	$R^2$	RPEE (%)	RMSE	$R^2$
$\mathbf{G}_1$	10.93	0.2615	0.8416	7.19	0.1889	0.9175	1.79	0.0315	0.9978
$\mathbf{G}_2$	16.02	0.3381	0.7356	12.39	0.2825	0.8155	3.99	0.0699	0.9884
$\mathbf{G}_3$	11.78	0.2607	0.8426	9.40	0.2184	0.8886	4.41	0.0779	0.9855

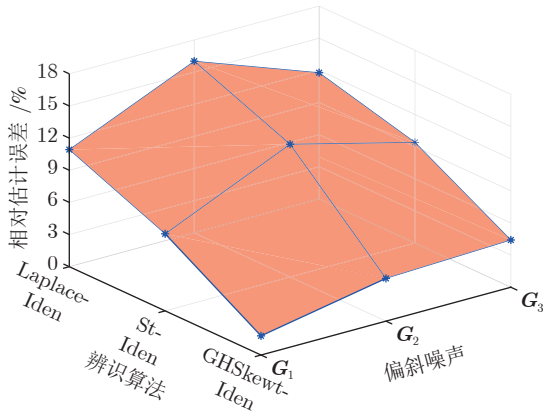


图 7 不同偏斜噪声下蒙特卡洛仿真实验的 RPEE 曲面  
Fig.7 The RPEE surface for the Monte Carlo with different levels of skewed noise

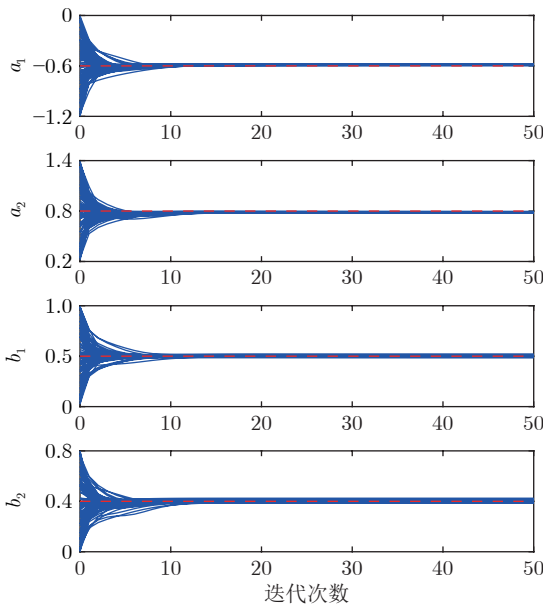


图 8 参数初始值不同时, 蒙特卡洛仿真实验的参数估计轨迹  
Fig.8 Trajectories of the parameter estimates for the Monte Carlo with different parameter initial values

表 4 参数初始值不同时蒙特卡洛仿真的均值和标准差  
Table 4 The mean and standard deviation of the Monte Carlo with different parameter initial values

模型参数	真实值	均值	标准差
$a_1$	-0.6000	-0.5860	0.0045
$a_2$	0.8000	0.7872	0.0033
$b_1$	0.5000	0.5028	0.0068
$b_2$	0.4000	0.4037	0.0080

外力  $F(t)$  为系统输入数据. 图 10 中, 输入数据纵坐标为质量块外力  $F$  (N), 质量块位置  $S(t)$  选取为

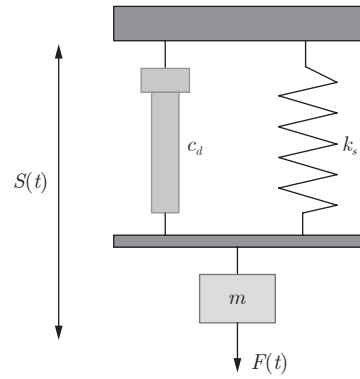


图 9 质量弹簧阻尼系统  
Fig.9 The mass-spring-damper system

系统的输出数据; 输出数据纵坐标为质量块位置  $S$  (m). 为了生成辨识数据, 将输入数据  $F(t)$  选取为幅值为 0.5 的随机二进制信号, 并用该信号激励质量弹簧阻尼系统得到输出数据  $S(t)$ . 为了测试本文 GHSkewt-Iden 算法对有偏噪声的有效性, 在输出数据上添加尺度、偏斜和自由度参数分别为  $[0.001, -0.2, 60]$  的偏斜噪声, 输出对比如图 11 所示, 其中蓝色实线表示真实输出, 红色虚线表示含偏斜噪声输出.

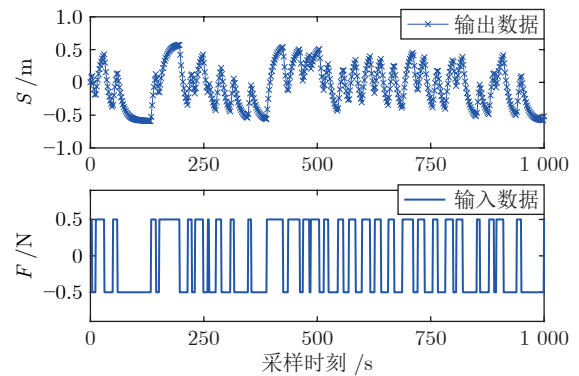


图 10 质量弹簧阻尼系统的输入输出数据  
Fig.10 The input and output of the mass-spring-damper system

在偏斜噪声条件下, 分别执行 GHSkewt-Iden、St-Iden 和 Laplace-Iden 算法, 可得到 3 个辨识模型. 利用图 10 的输入数据  $F(t)$ , 分别激励 3 种算法辨识得到的 3 个模型, 自我验证输出估计对比如图 12 所示. 在自我验证中, 计算得到的输出评价指标 RMSE 和  $R^2$  如表 5 所示. 由表 5 可知, 基于本文 GHSkewt-Iden 算法得到的模型输出评价指标 RMSE 值最小, 分别比 St-Iden 和 Laplace-Iden 算法小 5 倍和 6 倍左右; 基于本文 GHSkewt-Iden 算法得到的模型输出评价指标  $R^2$  值最大, 更趋近于 1. 上述辨识结果说明了, 基于本文 GHSkewt-Iden

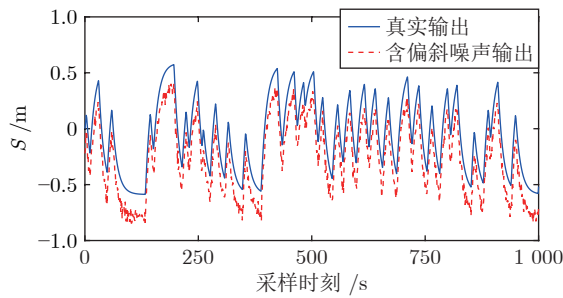


图 11 质量弹簧阻尼系统的真实输出和含偏斜噪声输出  
Fig. 11 The real output and output with skewed noise of the mass-spring-damper system

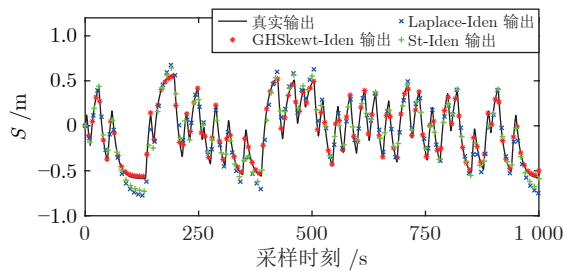


图 12 自我验证输出估计对比

Fig. 12 Output estimation comparison of self-verification

算法得到的模型输出更接近真实输出, 相比其他 2 种算法, 得到的模型更能精准地描述质量弹簧阻尼系统的动态特性.

表 5 自我验证的 RMSE 和  $R^2$

Table 5 The RMSE and  $R^2$  for self-verification

质量块位置	Laplace-Iden		St-Iden		GHSkewt-Iden	
	RMSE	$R^2$	RMSE	$R^2$	RMSE	$R^2$
$S$ (m)	0.0795	0.9373	0.0607	0.9635	0.0137	0.9981

为了进一步验证 3 个模型的泛化性能, 设计了一组输入和输出长度  $N = 2000$  的数据进行交叉验证. 交叉验证输出估计对比如图 13 所示, 交叉验证输出估计误差如图 14 所示, 交叉验证的 RMSE 和  $R^2$  如表 6 所示. 可以看出, 基于本文 GHSkewt-Iden 算法得到的模型泛化性能更强.

综上, 基于上述辨识结果可得出以下结论:

1) 由图 12 和表 5 可以看出, 在偏斜噪声存在条件下, 基于本文 GHSkewt-Iden 算法得到的模型输出更接近系统的真实输出数据, 该模型能够更加精确地描述质量弹簧阻尼系统的动态特性. 这证明了本文 GHSkewt-Iden 算法能够有效地处理偏斜噪声问题, 并且得到更加精确的模型参数值.

2) 由图 13、图 14 和表 6 可以看出, 在偏斜噪声条件下, 与 St-Iden 和 Laplace-Iden 算法得到的模型相比, 基于本文 GHSkewt-Iden 算法得到的模

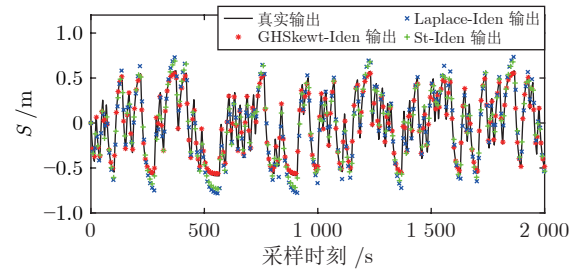


图 13 交叉验证输出估计对比

Fig. 13 Output estimation comparison of cross-verification

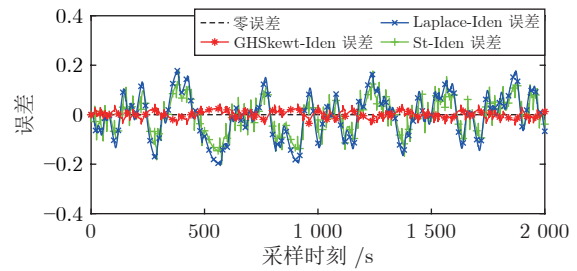


图 14 交叉验证输出估计误差

Fig. 14 Output estimation errors of cross-verification

表 6 交叉验证的 RMSE 和  $R^2$

Table 6 The RMSE and  $R^2$  for cross-verification

质量块位置	Laplace-Iden		St-Iden		GHSkewt-Iden	
	RMSE	$R^2$	RMSE	$R^2$	RMSE	$R^2$
$S$ (m)	0.0894	0.9345	0.0674	0.9627	0.0148	0.9982

型泛化性能更强. 这进一步验证了本文 GHSkewt-Iden 算法在质量弹簧阻尼系统上处理偏斜噪声的有效性, 可以更准确地拟合质量弹簧阻尼系统.

## 4 结束语

本文在非对称偏斜测量噪声条件下, 利用 GHSkewt 分布对输出噪声进行建模, 并提出线性系统鲁棒辨识 (GHSkewt-Iden 算法). 从数学上证明了标准学生氏  $t$  分布可看作是 GHSkewt 分布的一个特例, 验证了本文提出的算法适用范围更广、泛化性能更强. 实验结果表明: 1) 针对输出异常值问题, 本文 GHSkewt-Iden 算法的效果与基于学生氏  $t$ 、拉普拉斯分布的算法具有可对比性, 表明 GHSkewt 分布的重尾特性保障了算法的鲁棒性; 2) 针对非对称偏斜噪声问题, 本文 GHSkewt-Iden 算法效果明显优于基于学生氏  $t$ 、拉普拉斯分布算法, 表明 GHSkewt 分布的偏斜特性使其能够自适应地调整超参数, 以描述偏斜噪声的统计特性, 进而保障了算法对非对称偏斜噪声的有效性和稳定性.

综上, 针对偏斜噪声问题, 本文提出的 GHSkewt-Ide 鲁棒辨识算法与现有的鲁棒辨识算法相比, 适用范围更广、泛化性能更强. 在本文研究基础上, 未来的研究工作可集中在以下 2 个方面:

1) 进一步对算法的收敛性和初值选取进行理论研究;

2) 实际工业过程除受偏斜噪声、异常值等干扰外, 常具有时变特性且新数据以流的方式传输, 因此研究基于流数据的在线辨识算法尤为重要.

## References

- Kodamana H, Huang B, Ranjan R, Zhao Y J, Tan R, Sammaknejad N. Approaches to robust process identification: A review and tutorial of probabilistic methods. *Journal of Process Control*, 2018, **66**: 68–83
- Wang Guo-Qing, Yang Chun-Yu, Ma Lei, Dai Wei. Nonlinear Kalman filter based on Gaussian-generalized-hyperbolic mixing distribution. *Acta Automatica Sinica*, 2023, **49**(2): 448–460 (王国庆, 杨春雨, 马磊, 代伟. 基于高斯-广义双曲混合分布的非线性卡尔曼滤波. *自动化学报*, 2023, **49**(2): 448–460)
- Adeniran A A, El Ferik S. Modeling and identification of nonlinear systems: A review of the multimodel approach. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2017, **47**(7): 1149–1159
- Chen X, Zhao S Y, Liu F. Robust identification of linear ARX models with recursive EM algorithm based on student's t-distribution. *Journal of the Franklin Institute*, 2021, **358**(1): 1103–1121
- Zhang J H, Zhang F, Chen M, Ding R Q, Xu B, Zong H Z. Parameter identification of hydraulic manipulators considering physical feasibility and control stability. *IEEE Transactions on Industrial Electronics*, 2024, **71**(1): 718–728
- Yang X Q, Yin S, Kaynak O. Robust identification of LPV time-delay system with randomly missing measurements. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2018, **48**(12): 2198–2208
- Hou J, Su H, Yu C Q, Chen F W, Li P H. Bias-correction errors-in-variables Hammerstein model identification. *IEEE Transactions on Industrial Electronics*, 2023, **70**(7): 7268–7279
- Mohammad J M. Online system identification using fractional-order Hammerstein model with noise cancellation. *Nonlinear Dynamics*, 2023, **111**: 7911–7940
- Liu Xin. Identification of linear time-delay systems with unknown delay distributions in its value range. *Acta Automatica Sinica*, 2023, **49**(10): 2136–2144 (刘鑫. 时滞取值概率未知下的线性时滞系统辨识方法. *自动化学报*, 2023, **49**(10): 2136–2144)
- Liu X, Zhang T T, Liu X F. Robust identification method for LPV ARX systems and its application to a mechanical unit. *IEEE Access*, 2019, **7**: 164418–164428
- Liu X P, Yang X Q. Variational identification of linearly parameterized nonlinear state-space systems. *IEEE Transactions on Control Systems Technology*, 2023, **31**(4): 1844–1854
- Liu X, Yang X Q, Yin S. Nonlinear system identification with robust multiple model approach. *IEEE Transactions on Control Systems Technology*, 2020, **28**(6): 2728–2735
- Yang X Q, Lin X, Li Z. Multimodel approach to robust identification of multiple-input single-output nonlinear time-delay systems. *IEEE Transactions on Industrial Informatics*, 2020, **16**(4): 2413–2422
- Baldacchino T, Worden K, Rowson J. Robust nonlinear system identification: Bayesian mixture of experts using the t-distribution. *Mechanical Systems and Signal Processing*, 2017, **85**: 977–992
- Ferdowsi H, Jagannathan S, Zawodniok M. An online outlier identification and removal scheme for improving fault detection performance. *IEEE Transactions on Neural Networks and Learning Systems*, 2014, **25**(5): 908–919
- Pearson R K. Outliers in process modeling and identification. *IEEE Transactions on Control Systems Technology*, 2002, **10**(1): 55–63
- Chen X, Zhao S, Liu F, Tao C B. Laplace distribution based online identification of linear systems with robust recursive expectation-maximization algorithm. *IEEE Transactions on Industrial Informatics*, 2023, **19**(8): 9028–9036
- Fang Y, Jeong M K. Robust probabilistic multivariate calibration model. *Technometrics*, 2008, **50**(3): 305–316
- Jin X, Huang B. Robust identification of piecewise/switching autoregressive exogenous process. *AIChE Journal*, 2010, **56**(7): 1829–1844
- Yang X Q, Lu Y J, Yan Z B. Robust global identification of linear parameter varying systems with generalised expectation-maximisation algorithm. *IET Control Theory and Applications*, 2015, **9**(7): 1103–1110
- Yang X Q, Liu X, Yin S. Robust identification of nonlinear systems with missing observations: The case of state-space model structure. *IEEE Transactions on Industrial Informatics*, 2019, **15**(5): 2763–2774
- Liu X, Lou S C, Dai W. Further results on “System identification of nonlinear state-space models”. *Automatica*, 2023, (2): Article No. 110760
- Yang X Q, Liu X P, Xu C. Robust mixture probabilistic partial least squares model for soft sensing with multivariate Laplace distribution. *IEEE Transactions on Instrumentation and Measurement*, 2021, **70**: 1–9
- Waheed A, Cai L. Alternative design for optical incremental encoder measurement systems. In: Proceedings of the IEEE International Conference on Industrial Technology. Taipei, China: IEEE, 2016. 634–639
- Nurminen H, Ardeshiri T, Piche R, Gustafsson F. Skew-t filter and smoother with improved covariance matrix approximation. *IEEE Transactions on Signal Processing*, 2018, **66**(21): 5618–5633
- Aas K, Haff I H. The generalized hyperbolic skew student's t-distribution. *Journal of Financial Econometrics*, 2006, **4**(2): 275–309
- Jouchi N, Yasuhiro O. Stochastic volatility model with leverage and asymmetrically heavy-tailed error using GHSkew student's t-distribution. *Computational Statistics and Data Analysis*, 2012, **56**(11): 3690–3704
- Bai M M, Huang Y L, Chen B D, Zhang Y G. A novel robust Kalman filtering framework based on normal-skew mixture distribution. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022, **52**(11): 6789–6805
- Hasegawa T, Yamaguchi R, Kakuta M, Ando M, Songee J, Tokuda I, et al. Application of state-space model with skew-t measurement noise to blood test value prediction. *Applied Mathematical Modelling*, 2021, **100**: 365–378
- Liu X, Wang C, Dai W. Probability-based identification of Hammerstein systems with asymmetric noise characteristics. *IEEE Transactions on Instrumentation and Measurement*, 2024, **73**: 1–11
- Jia J T, Guo K X, Li W S, Xiang Y, Guo L. Composite filtering for UWB-based localization of quadrotor UAV with skewed measurements and uncertain dynamics. *IEEE Transactions on Instrumentation and Measurement*, 2022, **71**: 1–13
- Yu M, Yang X Q, Liu X P. LPV system identification with multiple-model approach based on shifted asymmetric Laplace distribution. *International Journal of Systems Science*, 2021, **52**(7): 1452–1465

- 33 Liu X P, Yang X Q. Identification of nonlinear state-space systems with skewed measurement noises. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2022, **69**(11): 4654–4662
- 34 Saha S. Noise robust online inference for linear dynamic systems. *IEEE Signal Processing Letters*, 2015, **22**(11): 1898–1902
- 35 Yang X Q, Lu Y J, Yan Z B. Robust global identification of linear parameter varying systems with generalised expectation-maximisation algorithm. *Journal of Financial Econometrics*, 2015, **9**(7): 1103–1110



**刘鑫** 中国矿业大学人工智能研究院副教授. 2019 年获得哈尔滨工业大学博士学位. 主要研究方向为系统辨识, 数据驱动的过程建模和软测量方法. E-mail: [15B904027@hit.edu.cn](mailto:15B904027@hit.edu.cn)

(**LIU Xin** Associate professor at the Artificial Intelligence Research Institute, China University of Mining and Technology. He received his Ph.D. degree from Harbin Institute of Technology in 2019. His research interest covers system identification, data-driven process modeling, and soft sensor methods.)



**陈强** 中国矿业大学信息与控制工程学院硕士研究生. 主要研究方向为系统辨识.

E-mail: [qiangchen@cumt.edu.cn](mailto:qiangchen@cumt.edu.cn)

(**CHEN Qiang** Master student at the School of Information and Control Engineering, China University

of Mining and Technology. His main research interest is system identification.)



**王兰豪** 中国矿业大学炼焦煤资源绿色开发全国重点实验室副教授. 主要研究方向为复杂工业过程的工艺参数检测、优化决策与智能控制. 本文通信作者.

E-mail: [wanglanhao888@163.com](mailto:wanglanhao888@163.com)

(**WANG Lan-Hao** Associate professor at National Key Laboratory for Green Development of Coking Coal Resources, China University of Mining and Technology. His research interest covers process parameter detection, optimal decision making, and intelligent control of complex industrial process. Corresponding author of this paper.)



**代伟** 中国矿业大学信息与控制工程学院、人工智能研究院教授. 主要研究方向为复杂工业过程建模、运行优化与控制.

E-mail: [weidai@cumt.edu.cn](mailto:weidai@cumt.edu.cn)

(**DAI Wei** Professor at the School of Information and Control Engineering and Artificial Intelligence Research Institute, China University of Mining and Technology. His research interest covers modeling, operational optimization and control for complex industrial process.)