Properties and Future of the Skew Kalman Filters

Lamia Alyami
Centre for Environmental Mathematics,
Faculty of Environment, Science and Economy,
University of Exeter, Penryn Campus,
Cornwall TR10 9FE, United Kingdom.
Email: la424@exeter.ac.uk, lmlasloom@nu.edu.sa

Saptarshi Das, *Member, IEEE*
Centre for Environmental Mathematics,
Faculty of Environment, Science and Economy,
University of Exeter, Penryn Campus,
Cornwall TR10 9FE, United Kingdom.
Email: S.Das3@exeter.ac.uk, saptarshi.das@ieee.org

***Abstract*—The Kalman filter (KF) and its variants are powerful numerical tools for estimating the unmeasured states of dynamical systems. However, traditional KFs assume Gaussian noise in measurements and processes, which may not always hold in practice. This paper reviews recent developments in non-Gaussian Kalman filters with non-zero skewness, which have received relatively little attention despite their potential benefits. This paper mainly focuses on skew Kalman filters (SKF), which replace the Gaussian assumption with the closed skew normal (CSN) distribution. This allows SKF to capture outliers in dynamical systems, resulting in improved performance and flexibility. Although there is limited literature on skew Kalman filters, this study provides an overview of their potential and motivation for use in a wide range of scientific applications.**

***Index Terms*—Kalman filtering, closed skew normal, outliers**

I. Introduction

Kalman filter is the heart of state estimation problems under noisy observations where the filter is named after R.E. Kalman published his famous paper in control theory [1]. The KF algorithm recursively estimates state variables in noisy measurements to track the evolution of dynamical systems over time, which is not a trivial task. This algorithm is an optimal estimator in the case of linear-Gaussian state-space models by minimising the mean-squared error (MMSE) of the state estimation. The KF algorithm is mathematically described as a predictor–corrector method since it updates the estimates once the next sequence of measurements are available to reduce the uncertainty in the covariance matrix in the prediction step. The KF algorithm and its variants are widely used in different applications due to its robustness, such as in signal processing [2], navigation [3], robotics [4], control systems [5] and infectious disease [6]. For more details of the KF algorithm theory and applications, please see [7] and [8]. The assumption of linearity and Gaussianity may be invalid where most of the phenomena in real world are nonlinear and non-Gaussian that encouraged researchers to invent several extensions of the KFs such as the more common variants like extended Kalman filter (EKF) in [9], iterated extended Kalman filter (IKF) [10] and unscented Kalman filter (UKF) [11].

Assuming the normality of the error terms in the KF, it can perform efficiently if the posterior states have a normality trend. However, Gaussian distribution can diverge if the state follows the non-Gaussian pattern, where the Gaussian posterior is not tractable. There are many approaches that can be used to approximate the KF with non-Gaussianity to improve the performance. Among the variants of KFs with non-Gaussianity we will focus on the KF with non-zero skewness approximated by an asymmetric distribution. The skewness is considered in the state space models assumptions to model the outliers in the presence of very skewed observations. The contemporary skew-normal distribution was first introduced by Azzalini in [12] by proposing the univariate skew-normal (SN) distribution and generalising it to the multivariate case (MSN) in [13] and [14]. The purpose of the SN distributions is to consider the skewness by adding a parameter for regulating it, where most of them are generalizations to the normal distribution. Further developments of skewness distributions exist in the literature and they have different formulas but equivalent parameterizations. However, we shall restrict this paper to focus on the class of skewed distributions called the closed skew normal (CSN) distribution, which is introduced in [15]. The pioneering study that suggested using skewness in the KF framework was reported in [16] and was introduced as a skew Kalman filter (SKF) in the dynamic linear case which is based on the CSN distribution. In [17], there is an emphasis on the fact that the CSN has most of the interesting properties of the normal distribution (e.g. closed under conditional, marginalization and summation). These properties allow the derivation of the recursive KF based on the CSN distribution. Consequently, relevant studies were followed up that implement the SKF in different approaches to derive a family of KFs which are based on the CSN distribution. In [18], the authors proposed the ensemble Kalman filter (EnKF) based on the CSN for nonlinear systems and [19] applied the skewness in the traditional UKF based on the CSN distribution to capture the third order moments of the state vector which achieved higher accuracy over the Gaussian UKF. There are various other asymmetric distributions that involve the KF for example in [20], the KF is based on scale mixtures of CSN distribution that contain scale mixtures of normal distributions as a special case and KF with skew-t distribution in [21] and [22]. A question may arise as to why we need to extend the KF with skewness distributions, whereas the particle filters (PFs) and their variants, such as the sequential Monte Carlo [23] family of algorithms, bootstrap filtering [24] and Monte Carlo filtering [25], were proposed to handle general distributions with high potential to estimate unmeasurable states in non-Gaussian and nonlinear systems. Although PFs offer several benefits as a generalized state estimation technique, a significant challenge associated with their use is the high computational cost. The performance of PFs depends on the number of particles, which can increase the computational costs and cause longer software running times, particularly for high-dimensional dynamical systems [26]. Rest of the paper is organized as follows. In section II, the CSN distribution and its properties are discussed. In section III, the classical KF is introduced followed by the skew KF in section IV. The conclusion and future trends are presented in section V.

II. The Closed Skew Normal Distribution

The CSN distribution is defined in [15] based on the conditional multivariate model that is given in [27]. This model assumed $E_1 \sim N_p(0,\Sigma)$ and $E_2 \sim N_q(0,\Delta)$ be independent normal random vectors given as:

$$\begin{aligned} W &= \mu + E_1, \\ Z &= -\nu + DE_1 + E_2. \end{aligned} \tag{1}$$

The joint distribution of W and Z is:

$$\begin{pmatrix} W \\ Z \end{pmatrix} \sim N_{q+p}\left[\begin{pmatrix} \mu \\ -\nu \end{pmatrix}, \begin{pmatrix} \Sigma & \Sigma D' \\ D\Sigma & \Delta + D\Sigma D' \end{pmatrix}\right]. \tag{2}$$

Now let us consider that we have a random vector X having the conditional distribution as $(W \mid Z \geq 0)$:

$$X = f(w \mid Z \geq 0) = \frac{f_W(w)}{P(Z \geq 0)} P(Z \geq 0 \mid W = w), \tag{3}$$

$$X = f(w \mid Z \geq 0) = C\phi_p(w;\mu,\Sigma)\Phi_q[D(w-\mu);\nu,\Delta], \tag{4}$$

$$C^{-1} = \Phi_q\left(0;\nu,\Delta + D\Sigma D'\right). \tag{5}$$

Here, $\phi_p(;\mu,\Sigma)$ is the probability density function (pdf) a p-dimensional normal distribution with mean μ and covariance matrix Σ and $\Phi_q(;\mu,\Sigma)$ is the cumulative distribution function (CDF) of the univariate standard normal distribution with q-dimension. Then X is a CSN distribution and parametrized as:

$$X \sim CSN_{p,q}(\mu,\Sigma,D,\nu,\Delta). \tag{6}$$

where, $\mu \in \mathbb{R}^p$, $\nu \in \mathbb{R}^q$, covariance $\Sigma \in \mathbb{R}^{p\times p}$, Δ is a covariance in $\mathbb{R}^{q\times q}$ and D is the shape parameter vector in $\mathbb{R}^{q\times p}$. We should note that the parameter μ is centralisation parameter, and Σ is a scale matrix and not have the same interpretation of the mean vector and the covariance matrix in the Gaussian distribution case. Amongst the other three parameters, D is the shape parameter which regulate the skewness of the distribution and if $D = 0$ the density (4) reduces to the multivariate normal distribution. Whereas the ν parameter secures the closure properties of the CSN under conditioning and the parameter Δ ensures the closure under marginalization. In the following, we will summarize the most important properties of the CSN distribution.

A. Properties of the CSN Distribution

Proposition 1: If $X \sim CSN_{p,q}(\mu,\Sigma,D,\nu,\Delta)$, then the moment generating function m.g.f of X is given by:

$$M_X(t) = \frac{\Phi_q(D\Sigma t;v,\Delta + D\Sigma D^T)}{\Phi_q(0;v,\Delta + D\Sigma D^T)} \exp\{t^T\mu + \frac{1}{2}t^T\Sigma t\}, t \in \mathbb{R}^p. \tag{7}$$

Proposition 2: Closure under linear transformation.
Let $X \sim CSN_{p,q}$ and A be a $r \times p$ matrix of rank p and $r < p$ then

$$AX \sim \mathrm{CSN}_{r,m}\left(\mu_{\mathrm{A}}, \Sigma_{\mathrm{A}}, D_A, v_A, \Delta_{\mathrm{A}}\right),$$

where,

$$\begin{aligned} \mu_{\mathrm{A}} = \mathrm{A}\mu, \quad \Sigma_{\mathrm{A}} = \mathrm{A}\Sigma\mathrm{A}^{\mathrm{T}}, \quad D_{\mathrm{A}} = D\Sigma\mathrm{A}^{\mathrm{T}}\Sigma_{\mathrm{A}}^{-1}, \nu_{\mathrm{A}} = \nu, \\ \text{and} \quad \Delta_A = \Delta + D\Sigma D^{\mathrm{T}} - D\Sigma A^{\mathrm{T}}\Sigma_A^{-1}A\Sigma D^{\mathrm{T}}. \end{aligned} \tag{8}$$

Proposition 3: Closure under marginalization.
Let $X \sim CSN_{p,q}(\mu,\Sigma,D,v,\Delta)$ which is partitioned as $X = \left[x_1^T x_2^T\right]^T$ where $x_1 \in \mathbb{R}^k$ and $x_2 \in \mathbb{R}^{n-k}$ then

$$\mathbf{x}_1 \sim \mathrm{CSN}_{n,m}\left(\mu_1, \Sigma_{11}, D^*, \mathbf{v}, \Delta^*\right),$$

with

$$\begin{aligned} D^* &= D_1 + D_2\Sigma_{21}\Sigma_{11}^{-1}, \quad \text{and} \\ \Delta^* &= \Delta + D_2\left(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}\right)D_2^T, \end{aligned} \tag{9}$$

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}, \quad \text{and } D = (D_1 D_2)$$

The conditional distribution of x_2 given x_1 is:

$$x_2 \mid x_1 \sim \mathrm{CSN}_{,} \text{ as the following}$$

$$\left(\mu_2 + \Sigma_{21}\Sigma_{11}^{-1}\left(x_1 - \mu_1\right), \Sigma_{22.1}, D_2, v - D^*\left(x_1 - \mu_1\right), \Delta\right) \tag{10}$$

where,

$$\Sigma_{22.1} = \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}.$$

The converse is also true.

Proposition 4: Closure under summation.
Let $X_p \sim \mathrm{CSN}_{p,q}\left(\mu_x, \Sigma_x, D, v, \Delta\right)$ and $Y_p \sim N\left(\mu_y, \Sigma_y\right)$, then the sum $Z_p = X_p + \mathbf{Y}_p$ follows the CSN distribution:

$$Z_n \sim \mathrm{CSN}_{n,m}\left(\mu_z, \Sigma_z, D_z, \nu_z, \Delta_z\right), \tag{11}$$

where,

$$\begin{aligned} \mu_z = \mu_x + \mu_y, \quad \Sigma_z = \Sigma_x + \Sigma_y, \quad D_z = D\Sigma_x\Sigma_z^{-1}, \\ \nu_z = \nu, \quad \Delta_z = \Delta + \left(D - D_z\right)\Sigma_x D^T. \end{aligned} \tag{12}$$

These favourable features above preserve stability and facilitate statistical inference problems with CSN distribution and for more details and proof for these propositions see [28]. Fig. 1 shows the effect of the shape parameter D on the asymmetry of the density with different values of the shape parameter $D = 10, 3, 0, -10$ respectively, with keeping other parameters fixed as $\mu = 0$, $\Sigma = 1.2$, $\nu = 0$, $\Delta = 0.5$. For the negative values of D, the distribution would be displayed on the opposite side of the vertical axis. Moreover, it is

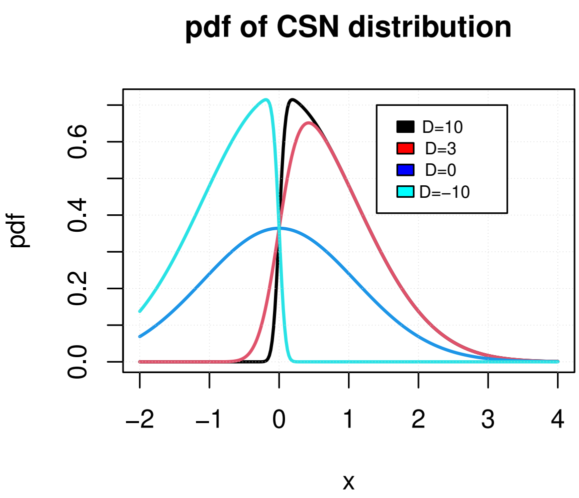


Fig. 1: PDF plot of the CSN distribution at different values of the skewness parameter D.

notable when $D = 0$ the density will reduce to the Gaussian distribution. In order to understand the the influence of the two additional parameters ν and Δ in the CSN distribution, we studied with different values and fixed other parameters as shown in Fig. 2. The shape of distribution is influenced by these two parameters. In the Fig. 2a with parameters values $\mu = 3$, $\Sigma = 9$, $D = 3$ and $\Delta = 1$. The negative values of the ν shift the curve to the negative area and tend to produce a flat curve or called platykurtic curve with $\Delta = -20$. The Fig. 2b with parameters values chosen as $\mu = 3$, $\Sigma = 9$, $D = 3$ and $\nu = 0$. There is a qualitative changes with varying values of Δ where the largest value of $\Delta = 20$ tends to have a bell shape with short tails in the right. These changes in the PDF curve adds flexibility to the CSN distribution by capturing complex noisy patterns in the observations which is one of the important properties as mentioned in [29]. However, technically the skewness parameter D is the only parameter that reserves the skewed pattern as shown in Fig. 1.

III. Classical Discrete Kalman Filter Revisited

In this section, we will introduce the standard derivation of the KF algorithm based on Gaussian distribution which is characterised by the first two moments without considering the higher order moments of non-Gaussian density. Let us consider the discrete-time dynamic system model as:

The process equation:

$$x_{t+1} = Fx_t + \eta_t, \quad \eta_t \sim (0, \Sigma_\eta).$$

The measurement equation:

$$y_t = Hx_t + \psi_t, \quad \psi_t \sim (0, \Sigma_\psi). \tag{13}$$

where, $x_t \in \mathbb{R}^n$ is the state vector of the system at time t, $y_t \in \mathbb{R}^m$ is the measurement vector, η_t and ψ_t are the process and measurement noises normally distributed with zero-mean and covariances Σ_η and Σ_ψ respectively. In addition, $F \in \mathbb{R}^{n\times n}$ is the state transition matrix and $H \in \mathbb{R}^{m\times n}$ is the measurement matrix. The KF algorithm has two stages: the prediction step and the correction or update step. The mean and covariance of the state x_t propagated with time in the KF equations as derived in (14) as:

$$\begin{aligned}
\hat{x}_0^+ &= E\left(x_0\right), \\
P_0^+ &= E\left[\left(x_0 - \hat{x}_0^+\right)\left(x_0 - \hat{x}_0^+\right)^T\right], \\
\hat{x}_t^- &= F\hat{x}_{t-1}^+, \\
P_t^- &= FP_{t-1}^+F^T + \Sigma_\eta, \\
K_t &= P_t^- H^T\left(HP_t^- H_t^T + \Sigma_\psi\right)^{-1}, \\
\hat{x}_t^+ &= \hat{x}_t^- + K_t\left(y_t - H\hat{x}_t^-\right), \\
P_t^+ &= \left(I - K_t H\right)P_t^-.
\end{aligned} \tag{14}$$

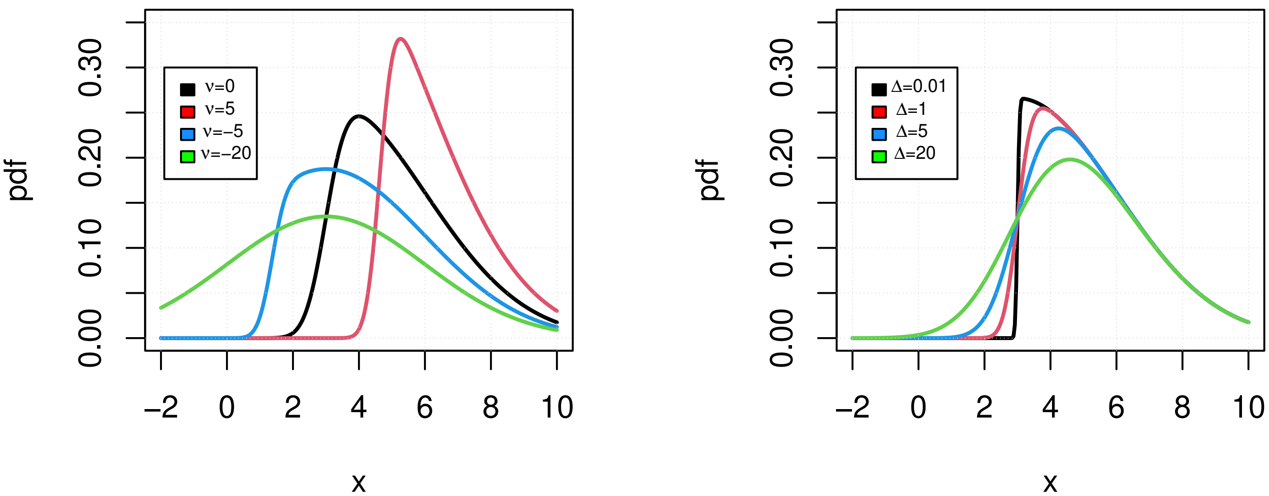
where, K_t is the Kalman gain, $\hat{x}_t^-$ and P_t^- are the estimate of the state and its covariance before the measurement y_t is processed where $\hat{x}_t^-$ is called a prior estimate. While $\hat{x}_t^+$ and P_t^+ are the estimate of the state and its covariance after giving the measurements where $\hat{x}_t^+$ is called the posterior estimate. For more details please see [30].

IV. Skew Kalman Filters and Variants

This section describes the main results of the paper reviewing existing works on SKFs. The previous sections were intended to provide a foundation for this section. The KF algorithm with skewness can be categorized based on its assumption of the CSN distribution - whether in the initial state vector or in the noise distribution. The SKF algorithm contains two parts the Gaussian part which is related to the smooth trend and depends on updating the mean μ and covariance Σ, the second part is the skewed part to capture the asymmetric pattern which depends on updating the skewness parameters D, ν and Δ. For the sake of completeness, we briefly review the differences and similarities in the prediction and update equations in the following cases.

A. Case 1: The State Vector x Following the CSN Distribution

The first proposed algorithm of the KF with CSN was in [16] which was for the linear state space model by assuming the initial state vector x_0 is distributed as $x_0 \sim CSN_{p,q}\left(\hat{\mu}_0, \hat{\Sigma}_0, \hat{D}_0, \hat{\nu}_0, \hat{\Delta}_0\right)$ whereas the process and measurement noises are Gaussian distributed with mean μ_η and covariance Σ_η, mean μ_ψ and covariance Σ_ψ respectively. The goal here is to capture the skewness of the state variables. The



(a) The CSN pdf with different values of ν.

(b) The CSN pdf with different values of Δ.

Fig. 2: The CSN distribution with different values of ν and Δ.

TABLE I: The list of recent references of SKF based on CSN distribution

Type of Skew Kalman Filters	Methods	Applications	Reference
SKF	linear system	experimental study	[16]
SKF	linear and nonlinear systems	petroleum reservoir simulations	[18]
Generalized of SKF	linear system	UK gas consumption	[31]
USKF	nonlinear systems	tracking velocity and position of a body	[19]
SKF	linear system	factor index of male and female anchovies	[32]

SKF equations in this case are:

$$\begin{aligned}
\hat{\mu}_t^- &= F_t \hat{\mu}_{t-1}^+ + \mu_\eta, \\
\hat{\Sigma}_t^- &= F_t \hat{\Sigma}_t^+ F_t^T + \Sigma_\eta, \\
\hat{\mu}_t^+ &= \hat{\mu}_t^- + K_t \left[y_t - H_t \hat{\mu}_t^+ \right] - \mu_\psi, \\
K_t &= \Sigma_t^- H_t^T \left(\Sigma_\psi + H_t \Sigma_t^- H_t^T \right)^{-1}, \\
\hat{\Sigma}_t^+ &= \hat{\Sigma}_t^- - K_t H_t \hat{\Sigma}_t^-, \\
\hat{D}_t &= \hat{D}_{t-1} \hat{\Sigma}_t^+ F_t^T \hat{\Sigma}_t^{-1}, \\
\hat{\nu}_t &= \hat{\nu}_{t-1}, \\
\hat{\Delta}_t &= \hat{\Delta}_{t-1} + \left(\hat{D}_{t-1} - \hat{D}_t F_t \right) \hat{\Sigma}_t^+ \hat{D}_{t-1}.
\end{aligned} \tag{15}$$

The equation (15) presents the general framework to integrate the skewness into the KF algorithm where the only difference is in the update step when propagating the skewness parameters D, ν and Δ. Despite the parameter ν is a constant but it is necessary to secure the conditional probability which is the fundamental situation in the KF algorithm. Furthermore, the summation of the CSN variable with a normal distribution variable is also the CSN distribution, for more details see lemma 2 in [16]. The proposition (4) helps to have a closed expression of the CSN distribution. It is noticeable that there is no big difference between the structure of the SKF and the Gaussian KF. The advantage of including the skewness in the initial state vector is that there is no change in the skewness dimension. Another study proposed in [19] shows that the CSN distribution in the state vector modifies the traditional unscented Kalman filter (UKF) [11] so that it can handle nonlinear systems based on the particles called the sigma points through unscented transformations to estimate the mean and covariance. The unscented SKF based on CSN distribution (CSN-UKF) is intended to capture the third-order moment (skewness) in the state vector where the process and measurement noise terms are assumed to be Gaussian distributed. The CSN-UKF as compared to the UKF has a significant performance with a lower mean squared error (MSE). However, they showed after a while the impact of skewness in the CSN-UKF will gradually converge to the Gaussian KF. The extension works for these assumptions are presented in [31] and [20] in greater detail.

B. Case 2: The Noise and State both Following the CSN Distribution

The approach of assuming the prior state vector and the process noise η being distributed the CSN distribution in both linear and nonlinear cases are proposed in [18]. Consequently, the recursive solution of the KF under these assumptions of the CSN distribution follows the CSN where the conjugacy is achieved under linear transformation and conditional distribution. The location μ and the scale Σ are identical to the algorithm in (15) but there are differences in skewness parameters within the time update step and this approach is computationally intensive. However, there is an increase in the skewness dimension where every time step has a different dimension and that is infeasible for practical implementations. To avoid exploring the skewness dimension some techniques might be useful such as re-fit every iteration with the CSN distribution as a prior state for the next step. This issue is discussed in [18] and by using numerical approximations as mentioned in [33]. This problem has been investigated earlier also in [16] by splitting up the state equation into a linear term and a skewed term. We are not giving the full derivation details of this approach and the reader should refer to the [18]. Further work addressed the skewness in the dynamic linear model in [32] where they introduced the filtering and smoothing scheme based on the generalized of the CSN distribution for observation noise since deep modifications of this topic are beyond the scope of the present paper but could be considered as future research. Table I provides a summary of the existing literature on SKFs related works including the difference in the methods and application areas.

V. Conclusions and Future Trends

We provide a brief survey of developments of sophisticated Kalman filtering algorithms for state estimation in the presence of non-zero skewness. Recent studies have shown that Skew Kalman filters offer better performance than the traditional KF. The SKF is mainly discussed based on the CSN distribution where the recursive filtering solution depends on the analytical properties of the CSN distribution. Recommendations for using different SKFs have also been discussed. Although in recent non-Gaussian signal processing, more focus has been given towards the particle filters since they do not impose assumption of any particular distribution in the recursive state estimation, rather they depend on the simulations. However, particle filters in higher dimensional state estimation problems are massively computationally expensive. On contrary, the SKFs assume that the noise, initial states or the state vectors can be parameterized as the CSN distributions which is much computationally cheaper than approximating the unknown posterior non-Gaussian distribution using many particles in the particle filters. The classical Kalman filters with Gaussian noise in all its variants still remain a powerful state estimation algorithm and have been successfully applied in different disciplines. However, further development of estimation and prediction problems needs to be developed with noise having non-zero skewness that may cover a wide range of real-world applications. For instance, the applications of SKF are limited and to the best of our knowledge, there is no previous work on applying SKF to estimate the COVID-19 pandemic states under various mechanistic epidemiological models. We hope this will inspire future researchers to focus on this particular filtering algorithm for state estimation of nonlinear and non-Gaussian problems. Another area of future research could be skew corrupted measurements along with model uncertainty where the parametric uncertainties of the model parameter posterior may also have a highly skewed distribution making such uncertain skew Kalman filtering problem more challenging.

Future methodological investigations on SKFs can be directed towards many areas e.g.

A. *Stability of SKFs*

The stability analysis of dynamical systems is very important and in KFs, it can be determined by looking at the eigenvalues of the Jacobian matrix of the state matrix. If we have outliers in the state variables, the Kalman gain can be unstable and that can not be solved by the Gaussian Kalman filters and in the case of SKF, it can be tackled using the stochastic stability analysis instead of the deterministic variant.

B. *Scalability of SKFs*

Scalability of the SKFs is the next important problem with an increasing number of estimated states for a fixed number of observations. As an example, the COVID-19 model that has been introduced in [6] from two measured states we had to estimate four hidden or unmeasurable states. By expanding this epidemiological model to be distributed geographically and with age-divided groups, it would need to replicate similar compartmental models where the number of unobserved states increases for fixed measurements (total death and infected counts) which makes such state estimation problems a high-dimensional one needing further investigation on scalability.

C. *SKFs with Higher Order Moments*

Further study can be directed towards other non-Gaussian Kalman filters assuming the noise is characterized by other specified higher-order moments like Kurtosis (4th central moment), super-skewness (5th central moment), super-flatness (6th central moment), hyper-skewness (7th central moment) and hyper-flatness parameters (8th central moment) [34], [35].

D. *Necessity and Limitations of SKFs*

The Laplace approximation is a way of approximating the shape of posterior distribution with an equivalent Gaussian distribution and most Kalman filters make this assumption for the noise distribution where accurately characterizing the noise is not trivial. However, for large deviations from normality, similar to the Laplace approximation, such filters suffer from the same drawbacks. On contrary, for fairly multivariate Gaussian noise, the Gaussian KFs will have higher efficiency over SKFs if the noise has a small skew parameter.

E. *SKFs for Nonlinear Systems*

Discrete-time nonlinear chaotic systems which are characterized by sensitivity to initial conditions exhibit challenges in state estimation. Most nonlinear KFs do not yield good results for chaotic systems and in the presence of skewed noise or multiplicative noise, such state estimation problems with nonlinear SKFs become even more complicated.

F. Comparison between SKFs and Particle Filters

In comparison with the non-parametric methods like particle filters (PFs) which are computationally expensive where the algorithm can run for a long time for high dimensional problems making it unreliable for long-term predictions. This is due to the fact that a small error in the initial state can be accumulated in later stages within PFs which usually do not happen in the standard KF as discussed in the limitations and drawbacks of PF in [36].

G. SKFs with Uncertain State Space Models

While quantifying the probability of failure for state space models with parametric uncertainty, the use of SKFs needs further investigation since the uncertainty on the model parameters may not be normal or maybe correlated in different states or even correlated with the noise within a multiplicative or more complex noise model. The above areas of SKFs need further theoretical and computational investigation for stability, computational feasibility, scalability and robustness of the state estimation algorithms using non-normal noise models.

Acknowledgment

LA thanks Najran University and Saudi Arabia Cultural Bureau in the UK, London for supporting this research. SD was partially supported by the ERDF Deep Digital Cornwall project number: 05R18P02820.

References

[1] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 83, no. 2, pp. 95–108, 1960.

[2] M. Roth, G. Hendeby, C. Fritsche, and F. Gustafsson, "The ensemble kalman filter: a signal processing perspective," *EURASIP Journal on Advances in Signal Processing*, vol. 2017, no. 1, pp. 1–16, 2017.

[3] G. Strang and K. Borre, *Linear algebra, geodesy, and GPS*. Siam, 1997.

[4] S. I. Roumeliotis and G. A. Bekey, "Bayesian estimation and kalman filtering: A unified framework for mobile robot localization," in *Proceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No. 00CH37065)*, vol. 3. IEEE, 2000, pp. 2985–2992.

[5] J. M. Kanieski, R. Cardoso, H. Pinheiro, and H. A. Gründling, "Kalman filter-based control system for power quality conditioning devices," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 11, pp. 5214–5227, 2012.

[6] L. Alyami and S. Das, "State estimation of the spread of covid-19 in saudi arabia using extended kalman filter," in *2022 Sensor Signal Processing for Defence Conference (SSPD)*. IEEE, 2022, pp. 1–5.

[7] A. Gelb, *Applied optimal estimation*. MIT press, 1974.

[8] P. Maybeck, "Stochastic niodels, estinurtion and contml: Volume i," 1979.

[9] A. Jazwinski, "Stochastic process and filtering theory, academic press," *A subsidiary of Harcourt Brace Jovanovich Publishers*, 1970.

[10] B. M. Bell and F. W. Cathey, "The iterated kalman filter update as a gauss-newton method," *IEEE Transactions on Automatic Control*, vol. 38, no. 2, pp. 294–297, 1993.

[11] S. J. Julier and J. K. Uhlmann, "New extension of the kalman filter to nonlinear systems," in *Signal processing, sensor fusion, and target recognition VI*, vol. 3068. International Society for Optics and Photonics, 1997, pp. 182–193.

[12] A. Azzalini, "A class of distributions which includes the normal ones," *Scandinavian journal of statistics*, pp. 171–178, 1985.

[13] A. Azzalini and A. D. Valle, "The multivariate skew-normal distribution," *Biometrika*, vol. 83, no. 4, pp. 715–726, 1996.

[14] A. Azzalini and A. Capitanio, "Statistical applications of the multivariate skew normal distribution," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 61, no. 3, pp. 579–602, 1999.

[15] G. González-Farıas, A. Domıńguez-Molina, and A. K. Gupta, "Additive properties of skew normal random vectors," *Journal of Statistical Planning and Inference*, vol. 126, no. 2, pp. 521–534, 2004.

[16] P. Naveau, M. G. Genton, and X. Shen, "A skewed kalman filter," *Journal of multivariate Analysis*, vol. 94, no. 2, pp. 382–400, 2005.

[17] M. G. Genton, *Skew-elliptical distributions and their applications: a journey beyond normality*. Chapman & Hall/CRC, 2004.

[18] J. Rezaie and J. Eidsvik, "Kalman filter variants in the closed skew normal setting," *Computational Statistics & Data Analysis*, vol. 75, pp. 1–14, 2014.

[19] ——, "A skewed unscented kalman filter," *International Journal of Control*, vol. 89, no. 12, pp. 2572–2583, 2016.

[20] H.-M. Kim, D. Ryu, B. K. Mallick, and M. G. Genton, "Mixtures of skewed kalman filters," *Journal of Multivariate Analysis*, vol. 123, pp. 228–251, 2014.

[21] H. Nurminen, T. Ardeshiri, R. Piché, and F. Gustafsson, "A nlos-robust toa positioning filter based on a skew-t measurement noise model," in *2015 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*. IEEE, 2015, pp. 1–7.

[22] ——, "Skew-t filter and smoother with improved covariance matrix approximation," *IEEE Transactions on Signal Processing*, vol. 66, no. 21, pp. 5618–5633, 2018.

[23] A. Doucet, N. De Freitas, and N. Gordon, "An introduction to sequential monte carlo methods," in *Sequential Monte Carlo methods in practice*. Springer, 2001, pp. 3–14.

[24] N. J. Gordon, D. J. Salmond, and A. F. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," in *IEE proceedings F (radar and signal processing)*, vol. 140, no. 2. IET, 1993, pp. 107–113.

[25] G. Kitagawa, "Monte carlo filter and smoother for non-gaussian nonlinear state space models," *Journal of computational and graphical statistics*, vol. 5, no. 1, pp. 1–25, 1996.

[26] Z. Chen *et al.*, "Bayesian filtering: From kalman filters to particle filters, and beyond," *Statistics*, vol. 182, no. 1, pp. 1–69, 2003.

[27] J. B. Copas and H. Li, "Inference for non-random samples," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 59, no. 1, pp. 55–95, 1997.

[28] G. González-Farías, J. Armando Domínguez-Molina, and A. K. Gupta, *The closed skew-normal distribution, in: M.G.Genton (Ed.),Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. CRC/Chapman Hall,Kondon, 2004.

[29] A. Azzalini, "An overview on the progeny of the skew-normal family—a personal perspective," *Journal of Multivariate Analysis*, vol. 188, p. 104851, 2022.

[30] D. Simon, *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[31] C. R. B. Cabral, C. Q. Da-Silva, and H. S. Migon, "A dynamic linear model with extended skew-normal for the initial distribution of the state parameter," *Computational Statistics & Data Analysis*, vol. 74, pp. 64–80, 2014.

[32] R. B. Arellano-Valle, J. E. Contreras-Reyes, F. O. L. Quintero, and A. Valdebenito, "A skew-normal dynamic linear model and bayesian forecasting," *Computational Statistics*, vol. 34, no. 3, pp. 1055–1085, 2019.

[33] O. Karimi, H. Omre, and M. Mohammadzadeh, "Bayesian closed-skew gaussian inversion of seismic avo data for elastic material properties," *Geophysics*, vol. 75, no. 1, pp. R1–R11, 2010.

[34] A. K. Majumdar, "Higher-order skewness and excess coefficients of some probability distributions applicable to optical propagation phenomena," *JOSA*, vol. 69, no. 1, pp. 199–202, 1979.

[35] R. Betchov and C. Lorenzen, "Phase relations in isotropic turbulence," *The Physics of Fluids*, vol. 17, no. 8, pp. 1503–1508, 1974.

[36] M. Jouin, R. Gouriveau, D. Hissel, M.-C. Péra, and N. Zerhouni, "Particle filter-based prognostics: Review, discussion and perspectives," *Mechanical Systems and Signal Processing*, vol. 72, pp. 2–31, 2016.