Robust Inference for State-Space Models with Skewed Measurement Noise

Henri Nurminen*, Tohid Ardeshiri†, Robert Piché*, and Fredrik Gustafsson†

henri.nurminen@tut.fi, tohid@isy.liu.se, robert.piche@tut.fi, fredrik@isy.liu.se

*Dept. of Automation Science and Engineering, Tampere University of Technology, Finland

†Division of Automatic Control, Linköping University, Sweden

Motivation

Heavy-tailed and skewed distributions arise e.g. in radio positioning, economics, biostatistics, and psychiatry.

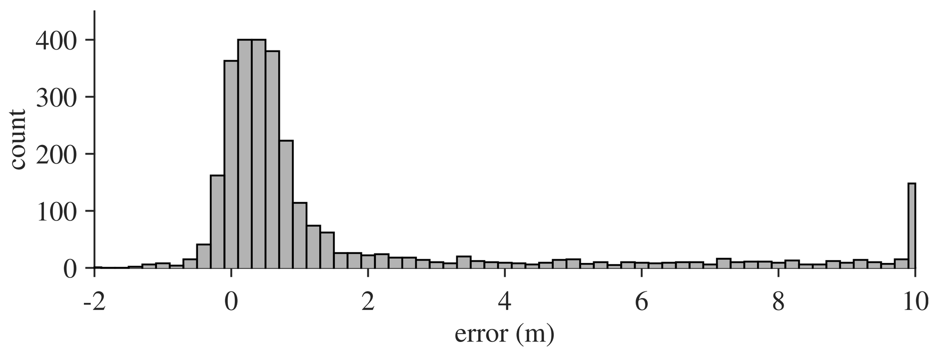

Figure 1: Non-line-of-sight causes skewness and outliers to TOA ranging error.

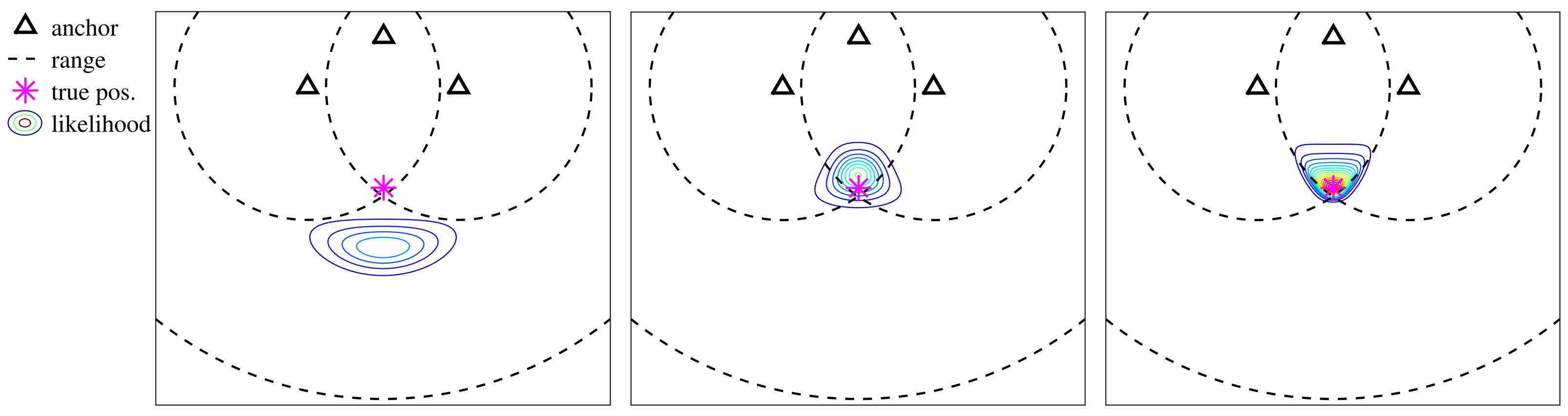

Figure 2: Student's t (middle) and skew-t (right) models accommodate an outlier, while Gaussian (left) gives a large estimation error.

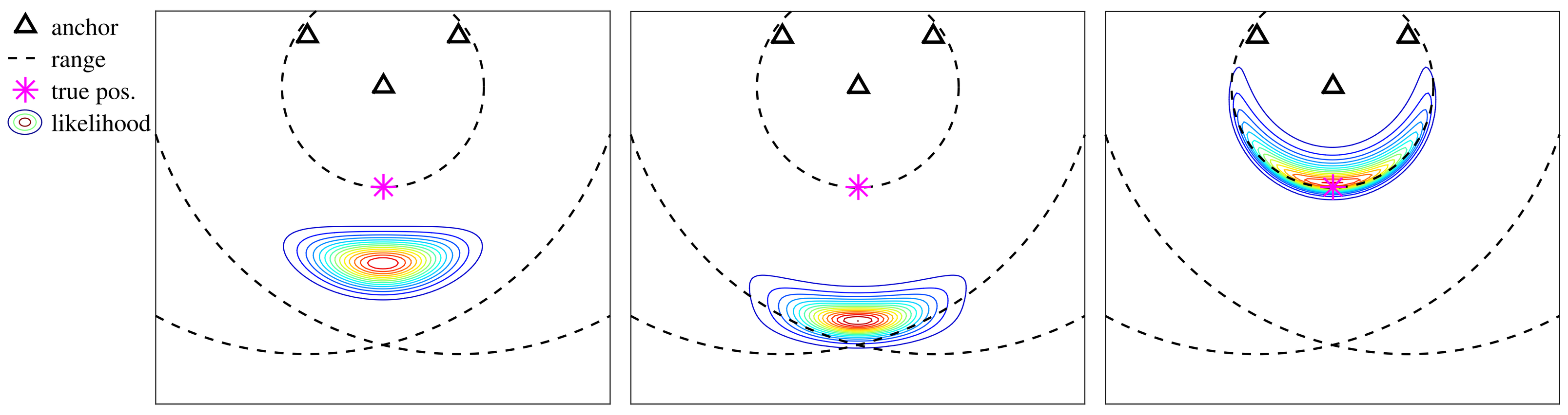

Figure 3: Skew t (right) uses the information that large negative outliers are improbable unlike Gaussian (left) and Student's t (middle).

Skew t-distribution

The skew t-distribution [5] is an extension of Student's t-distribution. $z \sim \mathrm{ST}(\mu, R, \Delta, \nu)$ has a hierarchical formulation as a Gaussian with random scaling and random bias with known sign:

$$z \mid u, \lambda \sim \mathrm{N}(\mu + \Delta u, \tfrac{1}{\lambda} R)$$
$$u \mid \lambda \sim \mathrm{N}_+(0, \tfrac{1}{\lambda} I)$$
$$\lambda \sim \mathrm{Gamma}(\tfrac{\nu}{2}, \tfrac{\nu}{2})$$

The parameters are
- μ: location
- Δ: skewness
- R: spread
- ν: degrees of freedom

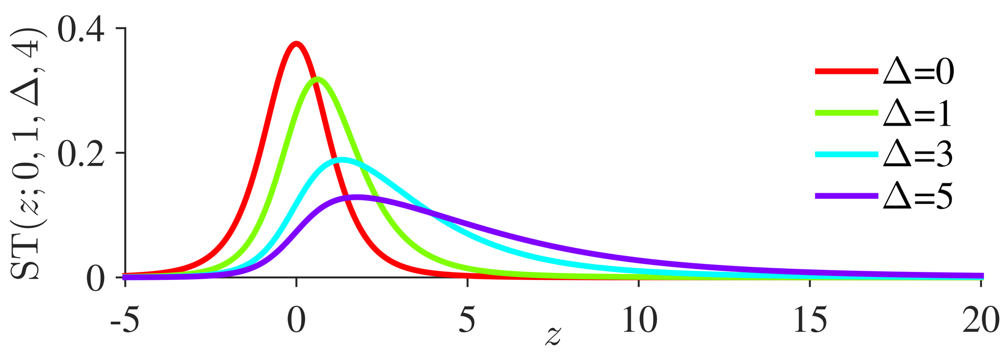

Figure 4: Skew-t densities with different Δs

Linear state-space model with skew-t measurement noise:

$$x_k = A x_{k-1} + w_{k-1}, \qquad w_{k-1} \sim \mathrm{N}(0, Q)$$
$$y_k = C x_k + e_k, \qquad e_k \sim \mathrm{ST}(\mu, R, \Delta, \nu).$$

Skew t variational Bayes filter [1]

The filtering posterior $p(x_k | y_{1:k})$ is not analytical, so we seek to approximate the posterior by

$$p(x_k, u_k, \lambda_k \mid y_{1:k}) \approx q(x_k)\, q(u_k)\, q(\lambda_k). \tag{1}$$

Variational Bayes (VB) gives optimal q functions in Kullback–Leibler sense. VB is an EM-type algorithm: update one variable at a time.

for $k = 1$ to K **do**
  Initialize $q(u_k)$ and $q(\lambda_k)$
  repeat
    Update $q(x_k) = \mathrm{N}(x_k; \cdot, \cdot)$ given $q(u_k)$ and $q(\lambda_k)$
    Update $q(u_k) = \mathrm{N}_+(u_k; \cdot, \cdot)$ given $q(\lambda_k)$ and $q(x_k)$
    Update $q(\lambda_k) = \mathrm{Gamma}(\lambda_k; \cdot, \cdot)$ given $q(x_k)$ and $q(u_k)$
  until Converged
  Predict $p(x_{k+1} \mid y_{1:k}) \approx \int p(x_{k+1} \mid x_k)\, q(x_k)\, \mathrm{d}x_k$
end for

GNSS & UWB positioning [1, 2]

Testing with simulated GNSS data and real UWB positioning data and comparison with t VB filter (TVBF) [4] and Kalman filter (KF).

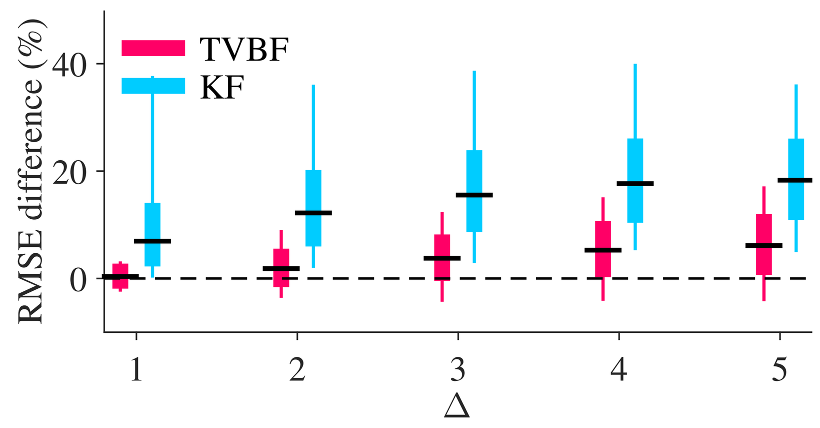

Figure 5: GNSS simulation

Table 1: UWB positioning

Filter	RMSE (m)	running time
EKF	1.36	1
STVBF	0.56	3

VB with recursive truncation [3]

The VB approximation (1) can show serious variance underestimation. We relax (1) so that x_k and u_k are not approximated as independent:

$$p(x_k, u_k, \lambda_k \mid y_{1:k}) \approx q(x_k, u_k)\, q(\lambda_k). \tag{2}$$

$q(x_k, u_k)$ is a truncated multivariate normal distribution, whose mean and covariance matrix can be approximated with the computationally light **recursive truncation** algorithm.

for $k = 1$ to K **do**
  Initialize $q(\lambda_k)$
  repeat
    Approximate $q(x_k, u_k) \approx \mathrm{N}([\begin{smallmatrix} x_k \\ u_k \end{smallmatrix}]; \cdot, \cdot)$ given $q(\lambda_k)$
    Update $q(\lambda_k) = \mathrm{Gamma}(\lambda_k; \cdot, \cdot)$ given $q(x_k, u_k)$
  until Converged
  Predict $p(x_{k+1} \mid y_{1:k}) \approx \int p(x_{k+1} \mid x_k)\, q(x_k)\, \mathrm{d}x_k$
end for

References

[1] Nurminen, Ardeshiri, Piché, Gustafsson, **Robust inference for state-space models with skewed measurement noise**, IEEE Signal Processing Letters, 2015.

[2] Nurminen, Ardeshiri, Piché, Gustafsson, **A NLOS-robust TOA positioning filter based on a skew-t measurement model**, International Conference on Indoor Positioning and Indoor Navigation (IPIN), 2015.

[3] Nurminen, Ardeshiri, Piché, Gustafsson, **Skew-t inference with improved covariance matrix approximation**, ArXiv, 2016.

[4] Piché, Särkkä, Hartikainen, **Recursive outlier-robust filtering and smoothing for nonlinear systems using the multivariate Student-t distribution**, IEEE International Workshop on Machine Learning and Signal Processing (MLSP), 2012.

[5] Sahu, Dey, Branco, **A new class of multivariate skew distributions with applications to Bayesian regression models**, Canadian Journal of Statistics, 2003.