Approximate Recursive Identification of Autoregressive Systems with Skewed Innovations

Henri Nurminen* and Tohid Ardeshiri†

henri.nurminen@tut.fi, ta417@cam.ac.uk

*Tampere University of Technology, Finland, †University of Cambridge, UK

Background

Heavy-tailed and skewed data series arise e.g. in radio positioning, financial time series, biostatistics, and psychiatry.

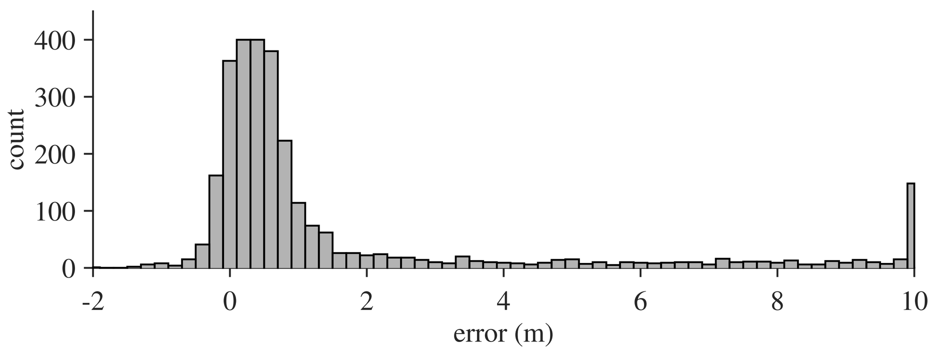

Figure 1: Non-line-of-sight causes skewness and outliers to TOA ranging error [2].

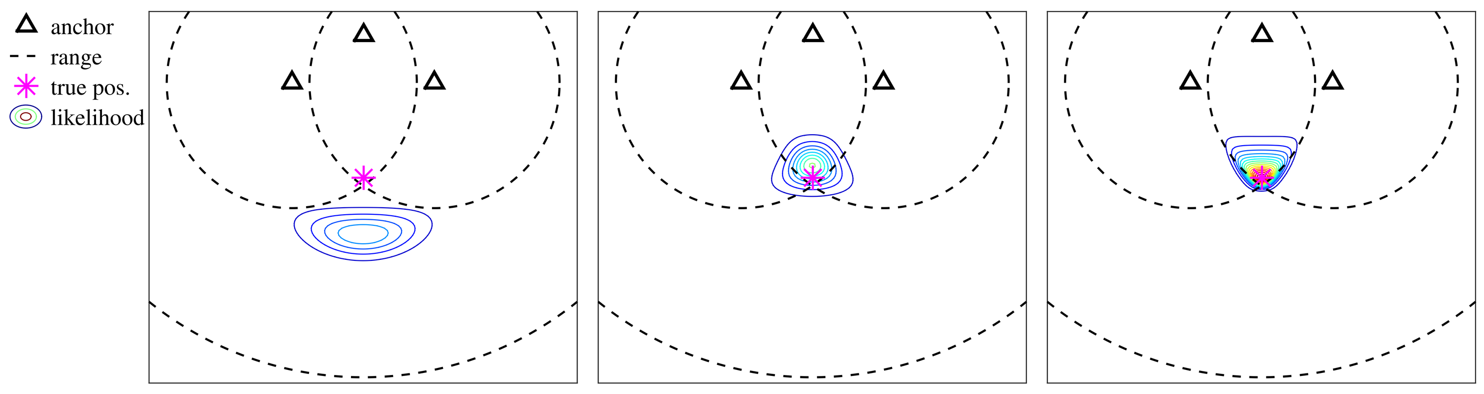

Figure 2: Student's t (middle) and skew-t (right) models accommodate an outlier, while Gaussian (left) gives a large estimation error [1].

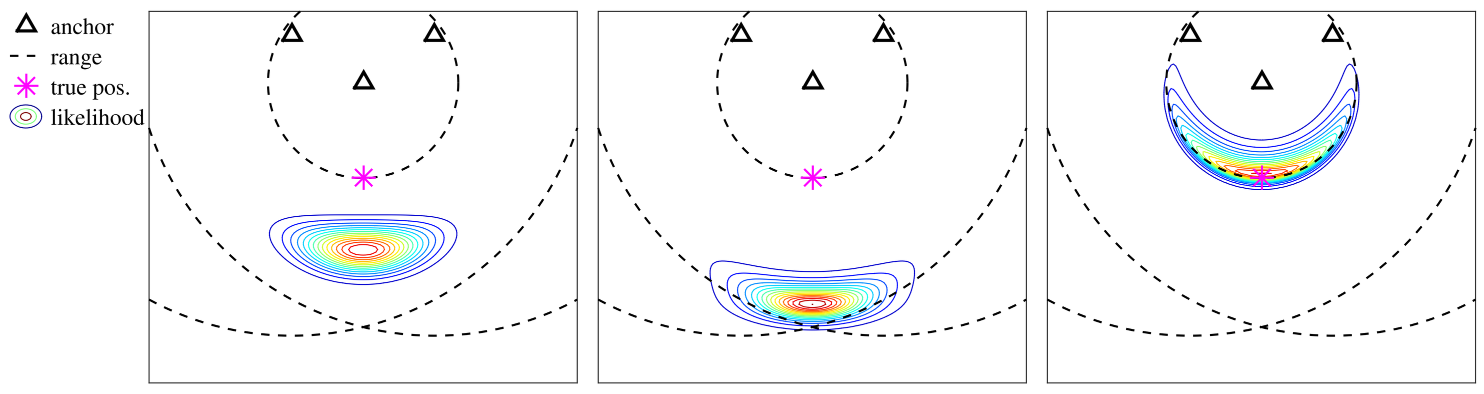

Figure 3: Skew t (right) uses the information that large negative outliers are improbable unlike Gaussian (left) and Student's t (middle) [1].

Skew normal and t-distributions

Extensions of Gaussian and Student-t-distribution. A multivariate skew-t variable $z \sim \mathrm{ST}(\mu, R, \Delta, \nu)$ [4, 5] has the hierarchical formulation

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

The parameters are

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

- $\lambda \equiv 1$ is skew normal.

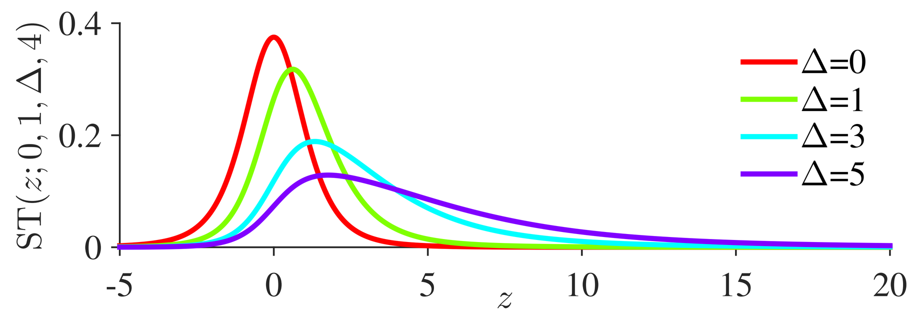

Figure 4: Skew-t densities with different Δs

Skew-t measurement update based on **variational Bayes** and **sequential truncation** approximations [1]:

repeat

$q(x_k, u_k) = \mathrm{N}_{\mathrm{trunc}}([\begin{smallmatrix} x_k \\ u_k \end{smallmatrix}]; \cdot, \cdot) \approx \mathrm{N}([\begin{smallmatrix} x_k \\ u_k \end{smallmatrix}]; \cdot, \cdot)$

$q(\lambda_k) = \mathrm{Gamma}(\lambda_k; \cdot, \cdot)$

until Converged

Recursive Skew-ARX identification

Assign the matrix-variate-normal–inverse-Wishart prior

$$
\begin{aligned}
p(R_k, \Delta_k) =& \mathrm{N}(\Delta_k; \Delta_{k|k-1}, R_k \otimes V_{k|k-1}) \\
& \times \mathrm{IW}(R_k; \Psi_{k|k-1}, \nu_{k|k-1})
\end{aligned}
$$

with a forgetting-factor type state transition and include in the variational iteration.

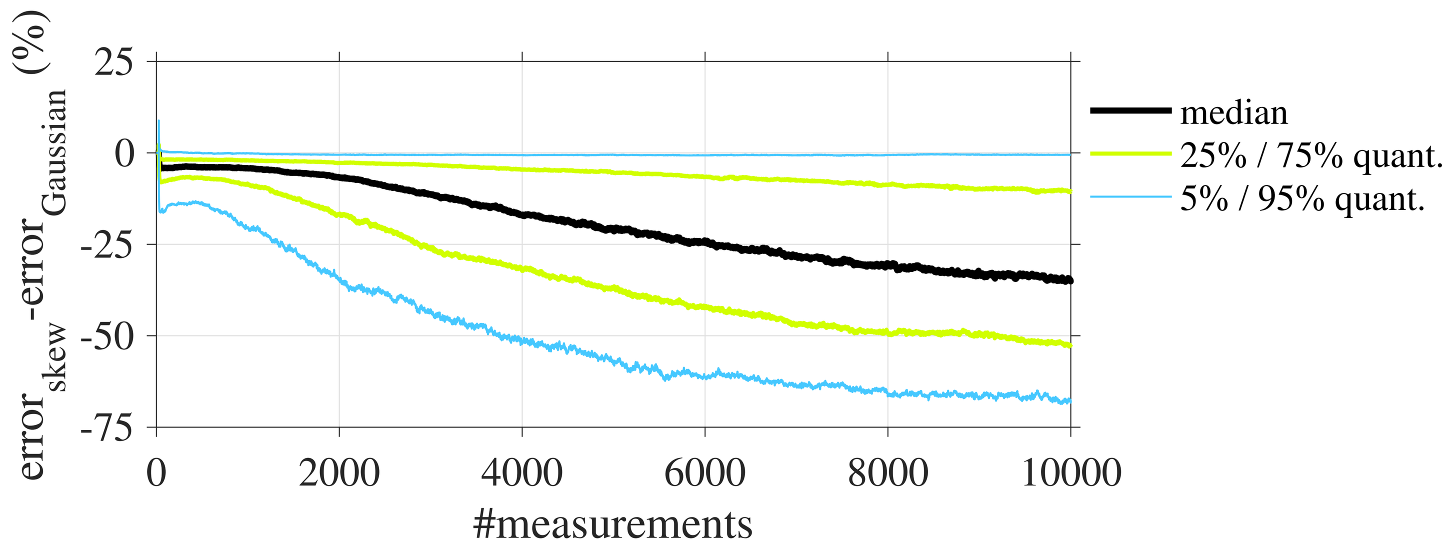

Figure 5: Simulation of AR(25) with skew-normal innovations. Skew-ARX outperforms the Gaussian algorithm in 95 % of the cases.

TOA positioning with skew-t filter

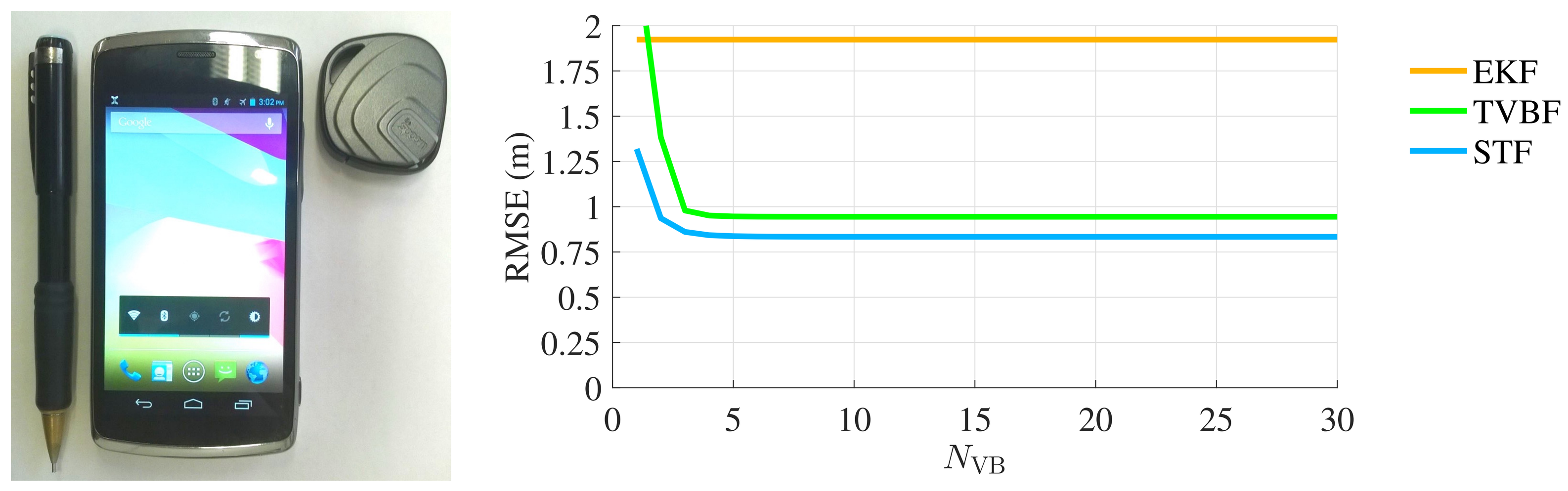

Figure 6: TOA-ranging based positioning using UWB or GNSS with skew-t filter and smoother[1, 3]. These extend Kalman filter and smoother.

Financial time series prediction

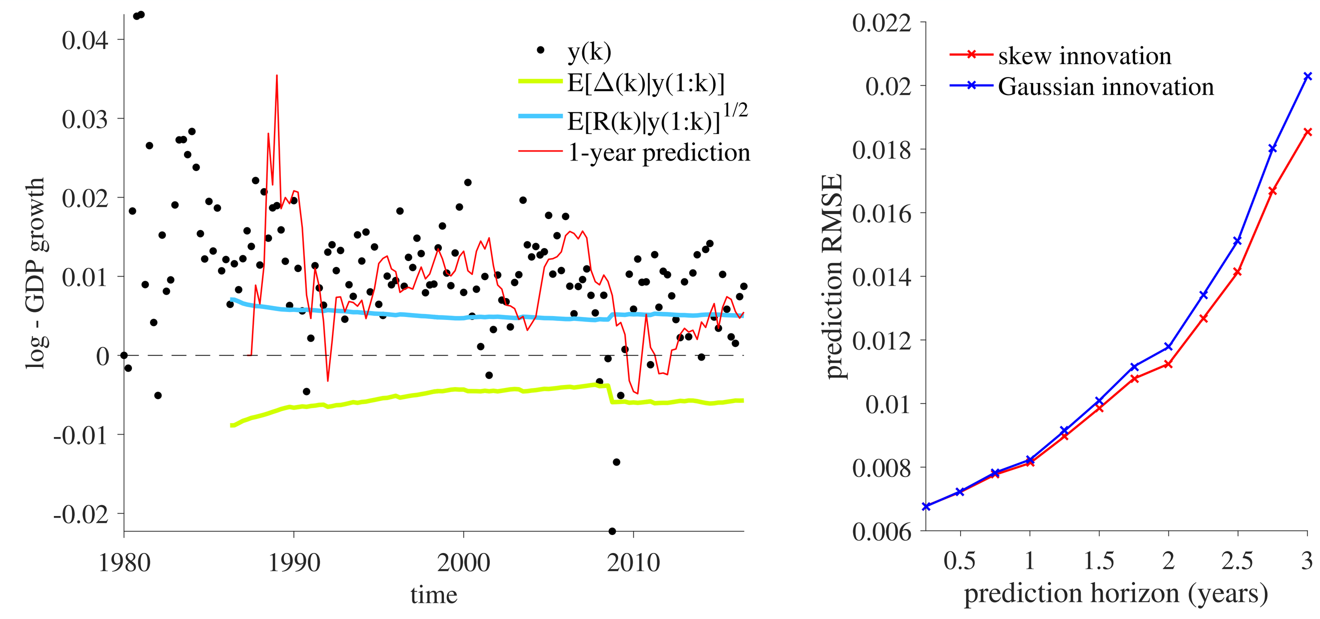

Figure 7: Quarterly US GDP prediction with Skew-AR(25)

- Skewed models are more flexible than Gaussian models
- Approximate state-space model inference and system identification
- VB approximation provides modest computational requirements and scalability
- Cramér–Rao lower bounds for filtering & smoothing [1]

References

[1] Nurminen, Ardeshiri, Piché, Gustafsson, **Skew-t filter and smoother with improved covariance matrix approximation**, http://arxiv.org/abs/1608.07435, 2016.

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

[3] 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.

[4] Azzalini, Dalla Valle, **The multivariate skew-normal distribution**, Biometrika, 1996.

[5] Lee, MacLachlan, **Finite mixtures of canonical fundamental skew t-distributions – the unification of the restricted and unrestricted skew t-mixture models**, Statistics and Computing, 2016.