

# Parameter identification using the skewed Kalman Filter

Katrin Runtemund, Gerhard Müller  
Chair for Structural Mechanics, TU München

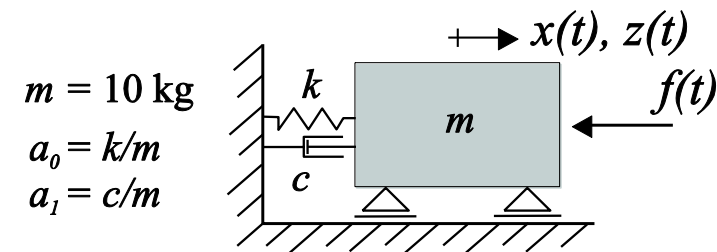
# Motivation and the investigated identification problem

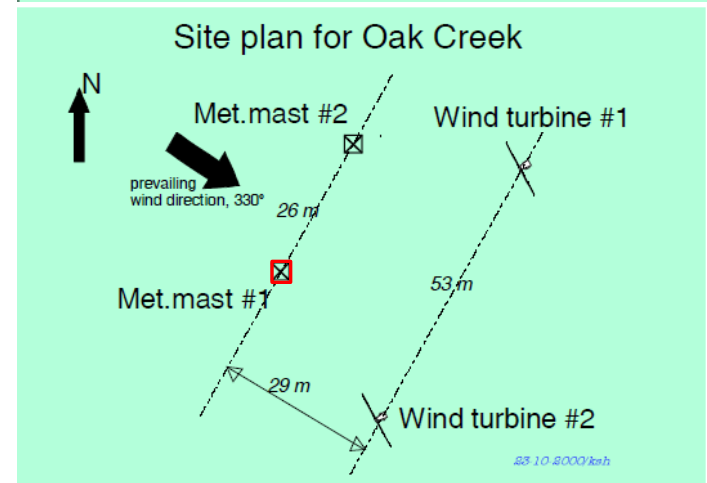
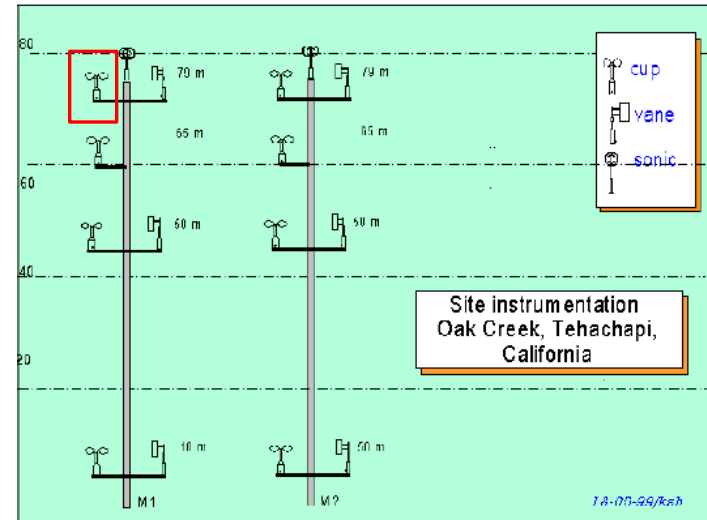
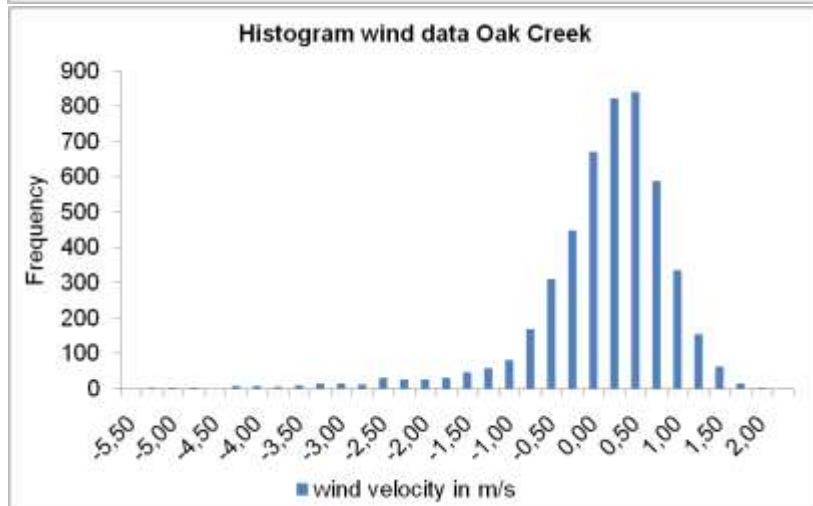
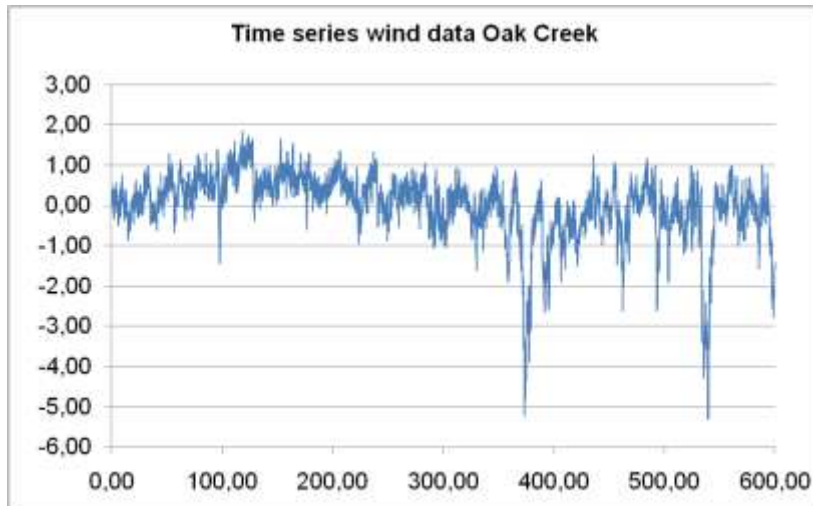
Problem: damped SDOF system excited by a **skewed process**

wanted: **estimation of the stiffness, damping parameter**  $a_0 = k/m$  and  $a_1 = c/m$

given:

- mathematical model of the system
- the statistical characteristics of the input force
  - ⇒ **wind data** was downloaded from the internet database: "Database of Wind Characteristics" located at DTU, Denmark, [www.winddata.com](http://www.winddata.com)) from the Oak Creek wind farm site
- the „true“ system response  $w(t)$  was simulated using the wind data





website: [www.winddata.com](http://www.winddata.com) of DTU

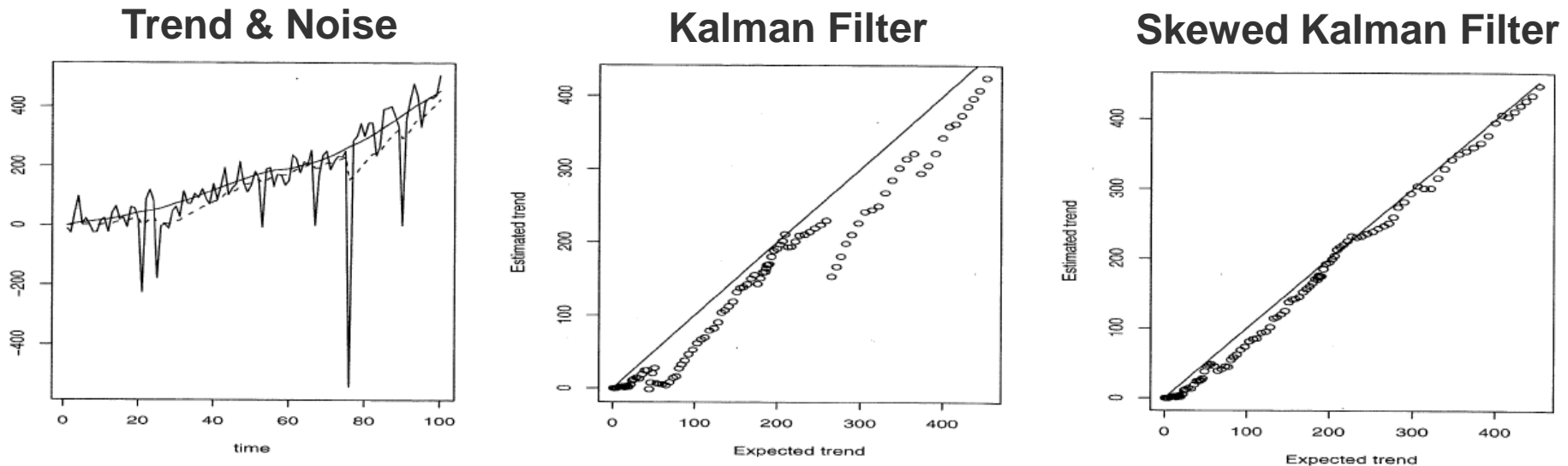
Wind speed measurement	mast 1; height 79 m; frequency 8 Hz; duration 600s
Statistics	mean 0 m/s; variance 0,69 m/s; skewness -1,99

# Overview

- Skewed Kalman Filter
- Extension of the skewed Kalman Filter for parameter identification problems
- Identification results of the application
- Conclusion

# Skewed Kalman Filter

# Skewed Kalman Filter (sKF)

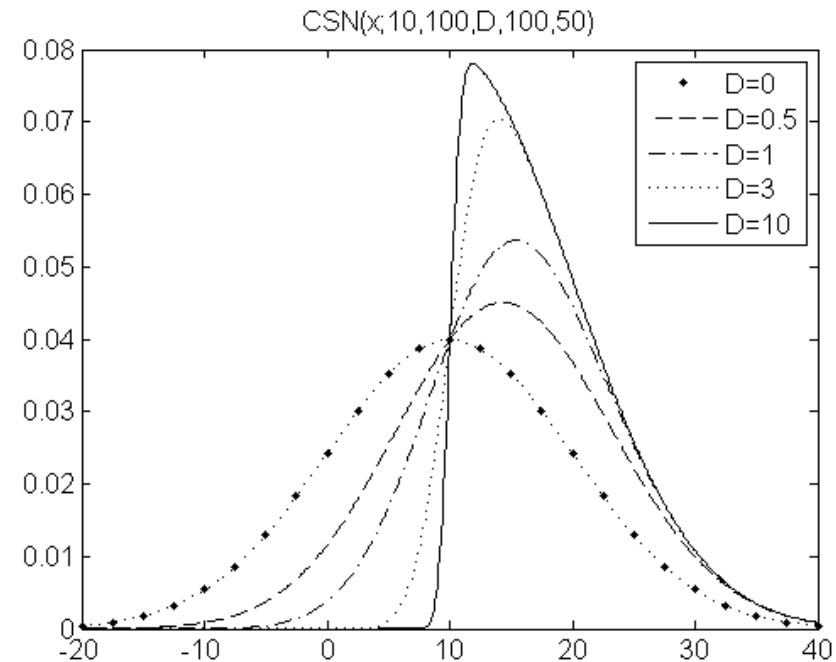


- introduced by P. Naveau, M. Genton and X. Shen  
Paper: “A skewed Kalman Filter”, *Journal of Multivariate Analysis*, 2004
- based on the **closed skew-normal distribution** (CSN)  
⇒ introduced by G. González-Farías, J.A. Domínguez-Molina, A. K. Gupta (2004)

Figure: M. G. Genton: Skew-Elliptical Distributions and their Applications – A Journal Beyond Normality

# Closed skew-normal distribution (CSN)

- family of distributions including the normal one, but with 3 additional parameters
- allows a continuous variation from a normal to a half-normal distribution
- closed under
  - ⇒ linear transformation
  - ⇒ marginalization
  - ⇒ conditioning
  - ⇒ summation\*



# Closed skew-normal distribution (CSN)

- CSN probability density function of a  $n$ -dimensional random vector  $\mathbf{X}$

$$X \sim CSN_{n,m}(\mu, \Sigma, D, \nu, \Delta) \quad p_{n,m}(\mathbf{x}) = \frac{\phi_n(\mathbf{x}; \mu, \Sigma) \Phi_m(\mathbf{D}(\mathbf{x} - \mu); \nu, \Delta)}{\Phi_m(\mathbf{0}; \nu, \Delta + \mathbf{D}\Sigma\mathbf{D}^T)}$$

mean values:  $\mu \in \mathbb{R}^n$  and  $\nu \in \mathbb{R}^m$

covariance matrices:  $\Sigma \in \mathbb{R}^{n \times n}$  and  $\Delta \in \mathbb{R}^{m \times m}$

arbitrary matrix  $D \in \mathbb{R}^{m \times n}$

$\phi_n(*; \eta, \Omega) = N(\eta, \Omega)$

$\Phi_n(*; \eta, \Omega)$ : cumulative distribution function

$C = \Phi_m(0; \nu, \Delta + D\Sigma D^T) = \text{const.}$

$D$  regulates skewness

continuous variation from the normal ( $D = 0$ ) to a half normal PDF

$\mu, \Sigma$  location and scale parameters

$\nu$  closure under conditioning

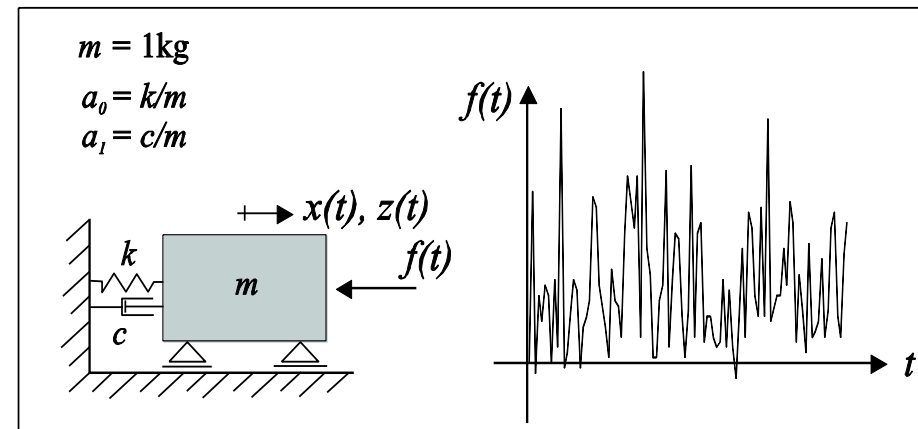
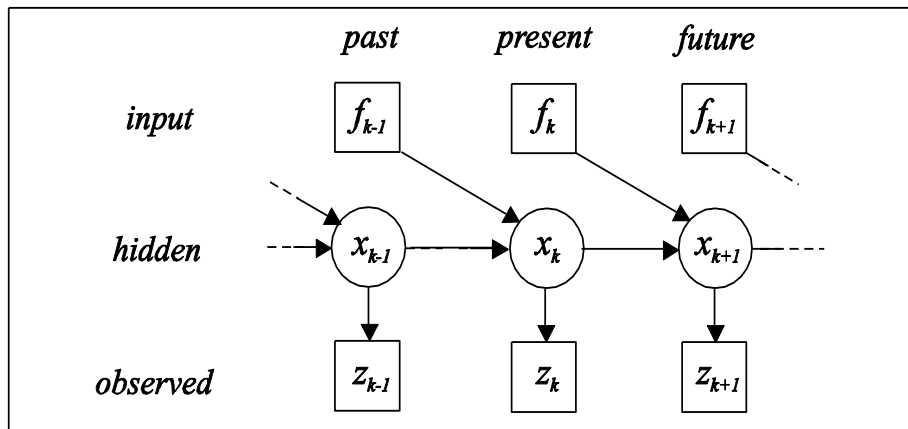
$\Delta$  closure under marginalization

$C$  closure under summation



# Standard Kalman Filter

- The standard KF is based on a linear state space model
  - $\Rightarrow$  linear system equation:  $\mathbf{x}_k = \mathbf{T}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{f}_k + \mathbf{S}\mathbf{w}_k \quad \mathbf{w}_k \sim \mathcal{N}(0, \Sigma_w)$
  - $\Rightarrow$  observation equation:  $\mathbf{z}_k = \mathbf{A}\mathbf{x}_k + \mathbf{v}_k \quad \mathbf{v}_k \sim \mathcal{N}(0, \Sigma_v)$
- the state follows a first order hidden Markov process represented as **Bayesian Dynamic Network**



# Implementation of skewness into the state space model

Alternative 1:

- Introducing the skewness in the system equation by modeling the input force as **additive CSN noise**:

$$\mathbf{f}_k \sim \text{CSN}(\mu, \Sigma, D, \nu, \Delta) \quad \longrightarrow \quad \mathbf{x}_k = \overbrace{\mathbf{T}\mathbf{x}_{k-1}}^{\text{skewed}} + \overbrace{\mathbf{B}\mathbf{f}_k}^{\text{skewed}} + \overbrace{\mathbf{S}\mathbf{w}_k}^{\text{noise}}$$

- problem: the sum of two CSN of dimension  $(n, m)$  is of dimension  **$(n, 2m)$**

$\Rightarrow$  rapid increase of the dimension of the matrices  $\mathbf{D}_k, \Delta_k$  at each time step

# Implementation of skewness into the state space model

Alternative 2:

- splitting up the state in a linear and a skewed part  $\mathbf{x}_k = [\mathbf{u}_k^T, \mathbf{s}_k^T]^T$

$$\mathbf{x}_k = \mathbf{T}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{f}_k + \mathbf{S}\mathbf{w}_k$$

$$\mathbf{z}_k = \mathbf{A}\mathbf{x}_k + \mathbf{v}_k$$



$$\mathbf{u}_k = \mathbf{T}\mathbf{u}_{k-1} + \mathbf{S}\mathbf{w}_k$$

$$\mathbf{z}_k = \mathbf{G}\mathbf{x}_k = \underbrace{\mathbf{A}\mathbf{u}_k}_{linear} + \underbrace{\mathbf{S}\mathbf{s}_k}_{skewed} + \underbrace{\mathbf{v}_k}_{noise}$$

- generation of the skewed process

$$\begin{bmatrix} \mathbf{X} \\ \mathbf{Y} \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} \boldsymbol{\mu} \\ \boldsymbol{\nu} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Sigma} & -\boldsymbol{\Sigma}\mathbf{D}^T \\ -\mathbf{D}\boldsymbol{\Sigma} & \Delta + \mathbf{D}\boldsymbol{\Sigma}\mathbf{D}^T \end{bmatrix} \right) \Rightarrow \mathbf{X} | \mathbf{Y} < \mathbf{D}\boldsymbol{\mu} \sim \text{CSN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathbf{D}, \boldsymbol{\nu}, \Delta)$$

# Implementation of skewness into the state space model

Alternative 2:

- splitting up the state in a linear and a skewed part  $\mathbf{x}_k = [\mathbf{u}_k^T, \mathbf{s}_k^T]^T$

$$\begin{array}{ccc}
 \mathbf{x}_k = \mathbf{T}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{f}_k + \mathbf{S}\mathbf{w}_k & \longrightarrow & \mathbf{u}_k = \mathbf{T}\mathbf{u}_{k-1} + \mathbf{S}\mathbf{w}_k \\
 \mathbf{z}_k = \mathbf{A}\mathbf{x}_k + \mathbf{v}_k & & \mathbf{z}_k = \mathbf{G}\mathbf{x}_k = \underbrace{\mathbf{A}\mathbf{u}_k}_{\text{linear}} + \underbrace{\mathbf{S}\mathbf{s}_k}_{\text{skewed}} + \underbrace{\mathbf{v}_k}_{\text{noise}}
 \end{array}$$

- generation of the skewed process

step 1: time-update  $\mathbf{y}_k = -\mathbf{L}_k\mathbf{y}_{k-1} + \eta_k$

step 2: determine the joint distribution

$$\begin{bmatrix} \mathbf{y}_k \\ \mathbf{y}_{k-1} \end{bmatrix} \sim \mathcal{N} \left( \begin{bmatrix} -\mathbf{L}_k\boldsymbol{\mu}_{k-1} + \boldsymbol{\mu}_{\eta,k} \\ \boldsymbol{\mu}_{k-1} \end{bmatrix}, \begin{bmatrix} \mathbf{L}_k\boldsymbol{\Sigma}_{k-1}\mathbf{L}_k^T + \boldsymbol{\Sigma}_{\eta,k} & -\boldsymbol{\Sigma}_{k-1}\mathbf{L}_k^T \\ -\mathbf{L}_k\boldsymbol{\Sigma}_{k-1} & \boldsymbol{\Sigma}_{k-1} \end{bmatrix} \right)$$

step 3: conditioning  $\mathbf{s}_k = \mathbf{y}_k \mid \mathbf{y}_{k-1} < \mathbf{D}_k\boldsymbol{\mu}_k \sim \text{CSN}(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \mathbf{D}, \boldsymbol{\nu}, \Delta)$

# Bayesian approach of the skewed Kalman filter

1. **Initialising** the filter using the prior PDF  $p(\mathbf{u}_{k-1}, \mathbf{y}_{k-1} | \mathbf{z}_{1:k-1})$
2. **Prediction**: time-update conditional on the observations up to time  $k-1$  using the system equations  $\hat{p}(\mathbf{u}_k, \mathbf{y}_k, \mathbf{y}_{k-1} | \mathbf{z}_{1:k-1})$
3. **Measurement update** after observing  $\mathbf{z}_k$ :  
 from the prediction error  $\mathbf{e}_k = \mathbf{z}_k - \mathbf{A}_k \bar{\boldsymbol{\mu}}_{u,k-1} - \mathbf{B}_k \mathbf{E}[\mathbf{s}_k | \mathbf{z}_{1:k-1}]$   
 likelihood  $p(\mathbf{z}_k | \mathbf{u}_k, \mathbf{y}_k, \mathbf{y}_{k-1}, \mathbf{z}_{1:k-1})$  is obtained using the observation equation
4. Applying the **Bayes' Rule** leads to the posterior PDF

$$p(\mathbf{u}_k, \mathbf{y}_k, \mathbf{y}_{k-1} | \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k | \mathbf{u}_k, \mathbf{y}_k, \mathbf{y}_{k-1}) p(\mathbf{u}_k, \mathbf{y}_k, \mathbf{y}_{k-1} | \mathbf{z}_{1:k-1})}{p(\mathbf{z}_k | \mathbf{z}_{1:k-1})}$$

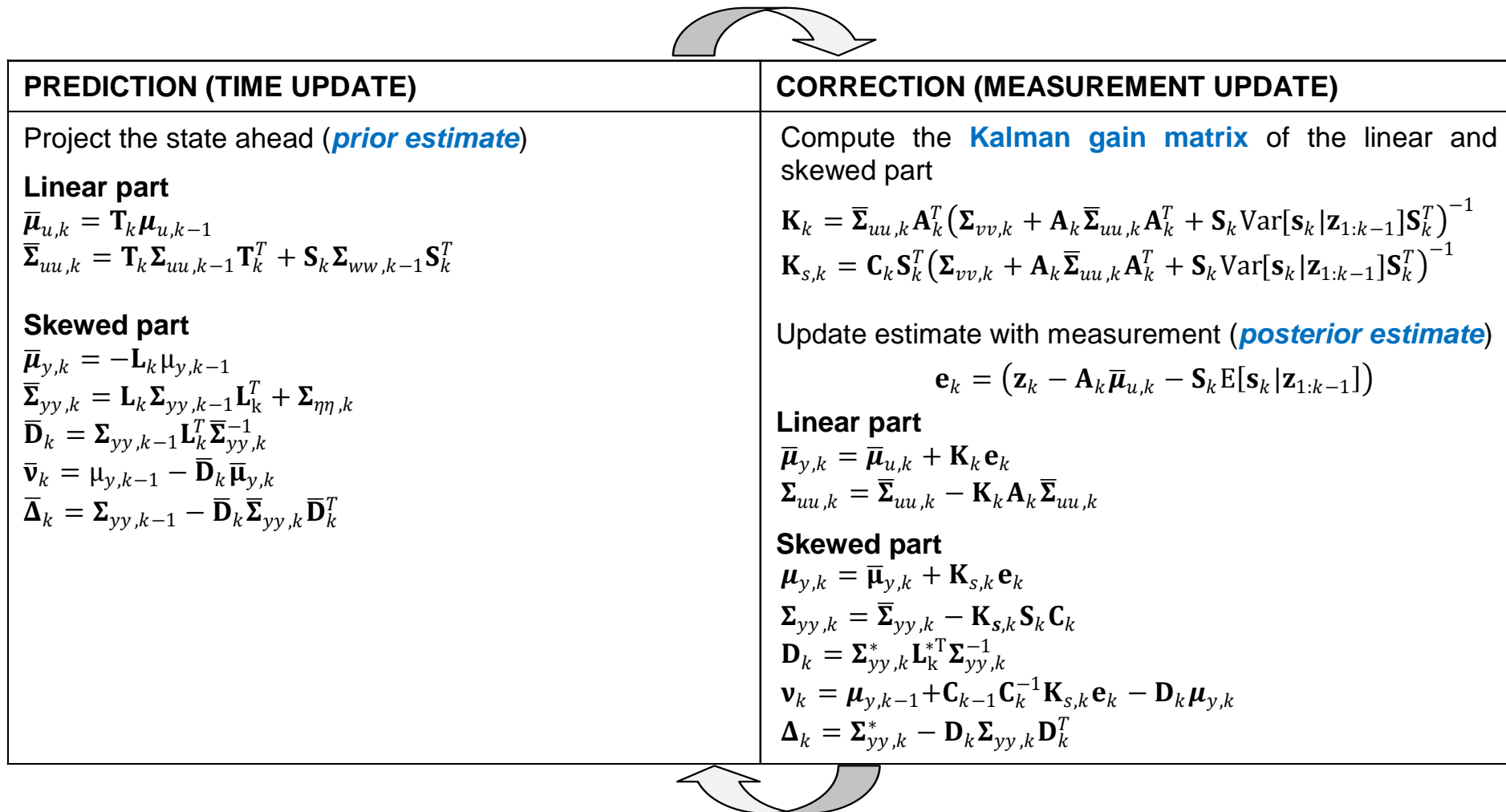
# Bayesian approach to the skewed Kalman filter

$$p(\mathbf{u}_k, \mathbf{y}_k, \mathbf{y}_{k-1} | \mathbf{z}_{1:k}) = \frac{p(\mathbf{z}_k | \mathbf{u}_k, \mathbf{y}_k, \mathbf{y}_{k-1}) p(\mathbf{u}_k, \mathbf{y}_k, \mathbf{y}_{k-1} | \mathbf{z}_{1:k-1})}{p(\mathbf{z}_k | \mathbf{z}_{1:k-1})}$$

5. **Marginalization** leads to the posterior PDF of the linear part  $p(\mathbf{u}_k | \mathbf{z}_{1:k})$
6. **Marginalization and Conditioning** with respect to  $\mathbf{y}_{k-1}$  leads to the posterior PDF of the skewed part  $p(\mathbf{s}_k | \mathbf{z}_{1:k})$

Advantages of the method:

- computationally efficient: the closure properties of the CSN allow to reduce the filter algorithm to the recursive calculation of the parameters
- At each time step a different skewness can be implemented by changing the matrix  $\mathbf{L}_k$  of the process  $\mathbf{y}_k = -\mathbf{L}_k \mathbf{y}_{k-1} + \eta_k$



PREDICTION (TIME UPDATE)	CORRECTION (MEASUREMENT UPDATE)
<p>Project the state ahead (<i>prior estimate</i>)</p> <p><b>Linear part</b></p> $\bar{\mu}_{u,k} = \mathbf{T}_k \mu_{u,k-1}$ $\bar{\Sigma}_{uu,k} = \mathbf{T}_k \Sigma_{uu,k-1} \mathbf{T}_k^T + \mathbf{S}_k \Sigma_{ww,k-1} \mathbf{S}_k^T$ <p><b>Skewed part</b></p> $\bar{\mu}_{y,k} = -\mathbf{L}_k \mu_{y,k-1}$ $\bar{\Sigma}_{yy,k} = \mathbf{L}_k \Sigma_{yy,k-1} \mathbf{L}_k^T + \Sigma_{\eta\eta,k}$ $\bar{\mathbf{D}}_k = \Sigma_{yy,k-1} \mathbf{L}_k^T \bar{\Sigma}_{yy,k}^{-1}$ $\bar{\mathbf{v}}_k = \mu_{y,k-1} - \bar{\mathbf{D}}_k \bar{\mu}_{y,k}$ $\bar{\Delta}_k = \Sigma_{yy,k-1} - \bar{\mathbf{D}}_k \bar{\Sigma}_{yy,k} \bar{\mathbf{D}}_k^T$	<p>Compute the <b>Kalman gain matrix</b> of the linear and skewed part</p> $\mathbf{K}_k = \bar{\Sigma}_{uu,k} \mathbf{A}_k^T (\Sigma_{vv,k} + \mathbf{A}_k \bar{\Sigma}_{uu,k} \mathbf{A}_k^T + \mathbf{S}_k \text{Var}[\mathbf{s}_k   \mathbf{z}_{1:k-1}] \mathbf{S}_k^T)^{-1}$ $\mathbf{K}_{s,k} = \mathbf{C}_k \mathbf{S}_k^T (\Sigma_{vv,k} + \mathbf{A}_k \bar{\Sigma}_{uu,k} \mathbf{A}_k^T + \mathbf{S}_k \text{Var}[\mathbf{s}_k   \mathbf{z}_{1:k-1}] \mathbf{S}_k^T)^{-1}$ <p>Update estimate with measurement (<i>posterior estimate</i>)</p> $\mathbf{e}_k = (\mathbf{z}_k - \mathbf{A}_k \bar{\mu}_{u,k} - \mathbf{S}_k \mathbb{E}[\mathbf{s}_k   \mathbf{z}_{1:k-1}])$ <p><b>Linear part</b></p> $\bar{\mu}_{y,k} = \bar{\mu}_{u,k} + \mathbf{K}_k \mathbf{e}_k$ $\Sigma_{uu,k} = \bar{\Sigma}_{uu,k} - \mathbf{K}_k \mathbf{A}_k \bar{\Sigma}_{uu,k}$ <p><b>Skewed part</b></p> $\mu_{y,k} = \bar{\mu}_{y,k} + \mathbf{K}_{s,k} \mathbf{e}_k$ $\Sigma_{yy,k} = \bar{\Sigma}_{yy,k} - \mathbf{K}_{s,k} \mathbf{S}_k \mathbf{C}_k$ $\mathbf{D}_k = \Sigma_{yy,k}^* \mathbf{L}_k^{*T} \Sigma_{yy,k}^{-1}$ $\mathbf{v}_k = \mu_{y,k-1} + \mathbf{C}_{k-1}^{-1} \mathbf{C}_k^{-1} \mathbf{K}_{s,k} \mathbf{e}_k - \mathbf{D}_k \mu_{y,k}$ $\Delta_k = \Sigma_{yy,k}^* - \mathbf{D}_k \Sigma_{yy,k} \mathbf{D}_k^T$

## Skewed Kalman Filter

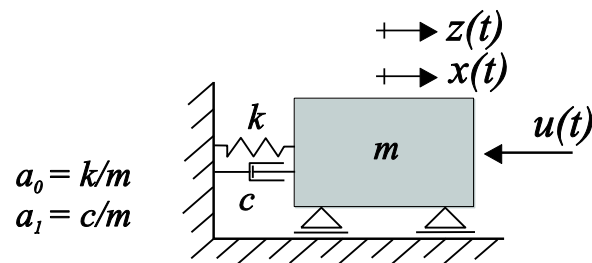
## Kalman Filter



# Extended skewed Kalman Filter (EsKF)

# Parameter identification using the Extended Kalman Filter (EKF)

Example: damped single degree of freedom system



wanted: **estimation of the stiffness, damping parameter  $a_0 = k/m$  and  $a_1 = c/m$**

$$\mathbf{u}_k = \mathbf{T} \mathbf{u}_{k-1} + \mathbf{S} \mathbf{w}_k$$

Problem: The transfer matrices  $\mathbf{T}_{k,k-1}$ ,  $\mathbf{B}_{k,k-1}$ ,  $\mathbf{S}_{k,k-1}$  depend on the unknown parameters:

$$T_k = e^{-\delta \Delta t} \begin{bmatrix} \frac{\delta}{\omega_d} \sin(\omega_d \Delta t) + \cos(\omega_d \Delta t) & \frac{1}{\omega_d} \sin(\omega_d \Delta t) \\ -\frac{\omega_0^2}{\omega_d} \sin(\omega_d \Delta t) & \cos(\omega_d \Delta t) - \frac{\delta}{\omega_d} \sin(\omega_d \Delta t) \end{bmatrix}$$

$$\begin{aligned} a_0 &= \frac{k}{m} = \omega^2 \\ a_1 &= \frac{c}{m} = 2\delta \end{aligned}$$

$$\omega_d = \sqrt{\omega^2 - \delta^2}$$

# Parameter identification using the Extended Kalman Filter (EKF)

step 1: Extending the state by the **parameter vector  $\mathbf{p}_k$**  to include the model parameters to be identified

$$x_{ext,k} = \begin{bmatrix} \mathbf{x}_k \\ \mathbf{p}_k \end{bmatrix} = \begin{bmatrix} w_k \\ v_k \\ a_{0,k} \\ a_{1,k} \end{bmatrix} \quad \longrightarrow \quad \begin{array}{l} \text{extended state space model} \\ x_{ext,k} = f(\mathbf{x}_{k-1}, \mathbf{p}_{k-1}, u_{k-1}, \mathbf{w}_{k-1}, \mathbf{w}_{p,k-1}) \\ z_{ext,k} = \mathbf{A}_{ext,k} \mathbf{x}_{ext,k} + \mathbf{v}_k \end{array}$$

step 2: Approximation of the nonlinear function  $f(\cdot)$  by a first order Taylor series near the current state estimate  $\hat{x}_{k-1}$

$$x_{ext,k} \approx f(\hat{x}_{0,k-1}) + \left. \frac{\partial f(x_{0,k-1})}{\partial x_{0i,k-1}} \right|_{\hat{x}_{0i,k-1}} (x_{0i,k-1} - \hat{x}_{0i,k-1})$$

$$x_{0,k-1} = \begin{bmatrix} \mathbf{x}_{k-1} \\ \mathbf{p}_{k-1} \\ u_{k-1} \\ \mathbf{w}_{k-1} \\ \mathbf{w}_{p,k-1} \end{bmatrix} \approx \begin{bmatrix} \hat{\mathbf{x}}_{k-1} \\ \hat{\mathbf{p}}_{k-1} \\ \hat{u}_{k-1} \\ \mathbf{0}_{2 \times 1} \\ \mathbf{0}_{2 \times 1} \end{bmatrix} = \hat{x}_{0,k-1}$$

# Parameter identification using the Extended Kalman Filter (EKF)

step 2: Approximation of the nonlinear function  $f(\cdot)$  by a first order Taylor series near the current state estimate  $\hat{\mathbf{x}}_{k-1}$

$$\mathbf{x}_{ext,k} \approx f(\hat{\mathbf{x}}_{0,k-1}) + \mathbf{T}_{ext,k} \underbrace{\begin{bmatrix} \mathbf{x}_{k-1} - \hat{\mathbf{x}}_{k-1} \\ \mathbf{p}_{k-1} - \hat{\mathbf{p}}_{p,k-1} \end{bmatrix}}_{\varepsilon_{ext,\hat{\mathbf{x}},k-1}} + \mathbf{B}_{ext,k} \underbrace{(\mathbf{u}_{k-1} - \hat{\mathbf{u}}_{k-1})}_{\varepsilon_{ext,u,k-1}} + \mathbf{S}_{ext,k} \underbrace{\begin{bmatrix} \mathbf{w}_{k-1} \\ \mathbf{w}_{p,k-1} \end{bmatrix}}_{\varepsilon_{ext,w,k-1}}$$

$$\mathbf{T}_{ext,k} = \begin{bmatrix} \mathbf{T}_k(\hat{\mathbf{p}}_{k-1}) & \mathbf{J}_{f,k}(\hat{\mathbf{p}}_{k-1}) \\ 0_{2 \times 2} & \mathbf{I}_{2 \times 2} \end{bmatrix} \quad \text{with} \quad \mathbf{J}_{f,k}(\hat{\mathbf{p}}_{k-1}) = \left. \frac{\partial f(\mathbf{x}_{0,k-1})}{\partial \mathbf{p}_{k-1}} \right|_{\hat{\mathbf{x}}_{0i,k-1}}$$

$$\mathbf{B}_{ext,k} = \begin{bmatrix} \mathbf{B}_k(\hat{\mathbf{p}}_{k-1}) \\ 0_{2 \times 1} \end{bmatrix}$$

$$\mathbf{S}_{ext,k} = \begin{bmatrix} \mathbf{S}_k(\hat{\mathbf{p}}_{k-1}) & \mathbf{J}_{f,k}(\mathbf{w}_{p,k-1}) \\ 0_{2 \times 1} & \mathbf{I}_{2 \times 2} \end{bmatrix} \quad \text{with} \quad \mathbf{J}_{f,k}(\mathbf{w}_{p,k-1}) = \left. \frac{\partial f(\mathbf{x}_{0,k-1})}{\partial \mathbf{w}_{p,k-1}} \right|_{\hat{\mathbf{x}}_{0i,k-1}}$$

Jacobian Matrix

$$\mathbf{J}_f(\mathbf{p}_k) = \begin{bmatrix} \frac{\partial w}{\partial a_0} & \frac{\partial v}{\partial a_0} \\ \frac{\partial w}{\partial a_1} & \frac{\partial v}{\partial a_1} \end{bmatrix}$$

# Parameter identification using the Extended skewed Kalman Filter (EsKF)

- in case of the skewed state space model, the linearization is done around the current state of the linear part

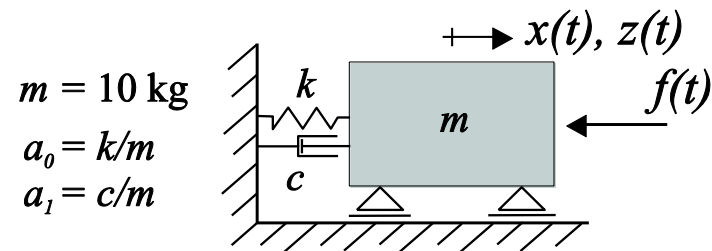
linear part:  $\mathbf{u}_{ext,k} = \mathbf{T}_{ext,k-1} \mathbf{u}_{ext,k-1} + \mathbf{S}_{ext,k-1} \mathbf{w}_{ext,k}$

skewed part:  $\mathbf{y}_k = -\mathbf{L}_k \mathbf{y}_{k-1} + \eta_k \Rightarrow \mathbf{s}_k = \mathbf{y}_k \mid \mathbf{y}_{k-1} < \mathbf{D}_k \mu_k$

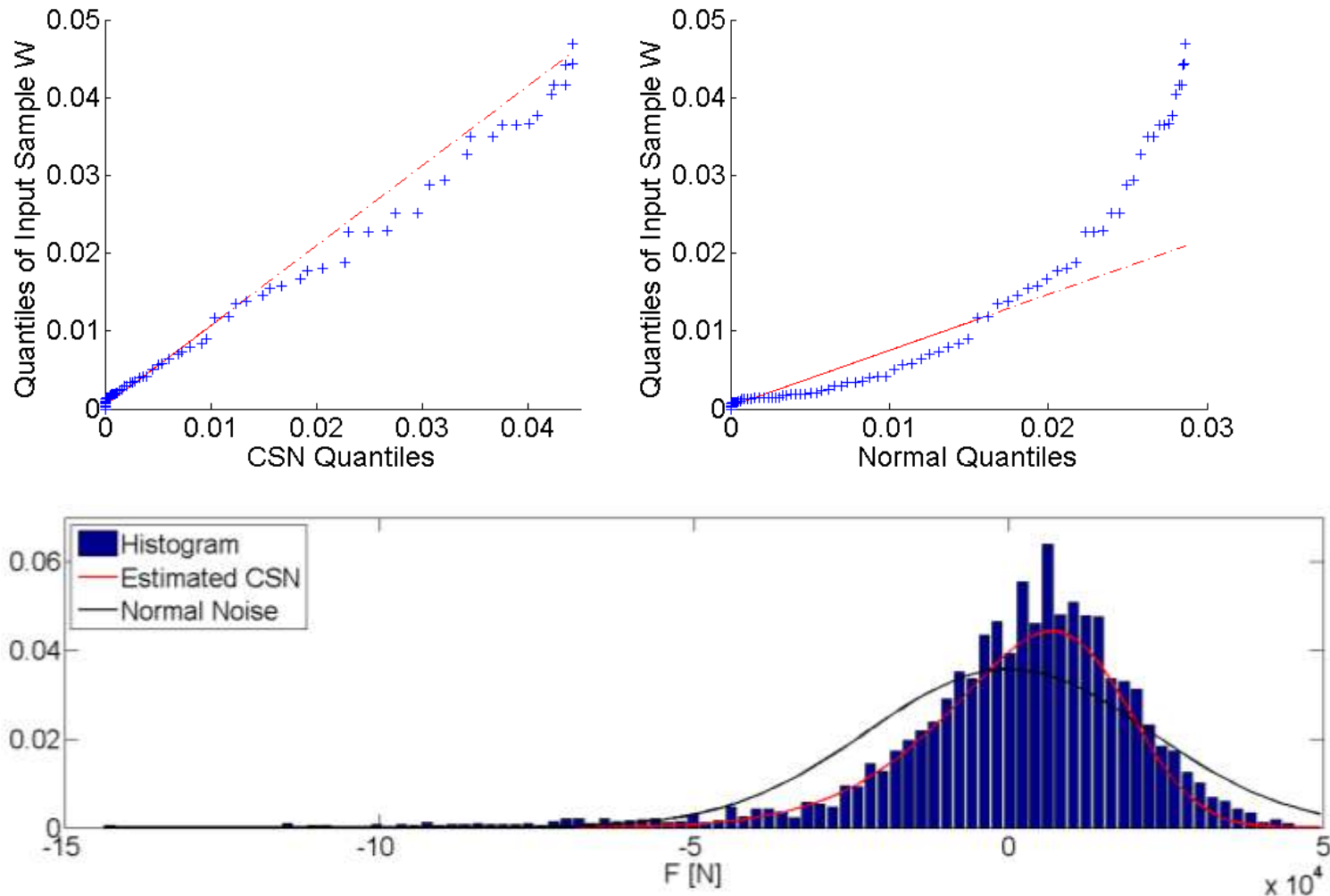
$$\mathbf{z}_k = \mathbf{G} \mathbf{x}_k = \mathbf{A}_{ext} \mathbf{u}_k + \mathbf{A}_{ext} \mathbf{S}_{ext,k-1} \mathbf{s}_k + \mathbf{v}_k \quad \mathbf{A}_{ext} = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$$

- the linearized extended state space model can now be investigated using the introduced skewed Kalman filter algorithm where  
 $\Rightarrow$  the time-variant transfer matrices  $\mathbf{T}_k$ ,  $\mathbf{S}_k$  are replaced by the extended matrices  $\mathbf{T}_{ext,k}$ ,  $\mathbf{S}_{ext,k}$

## Identification results

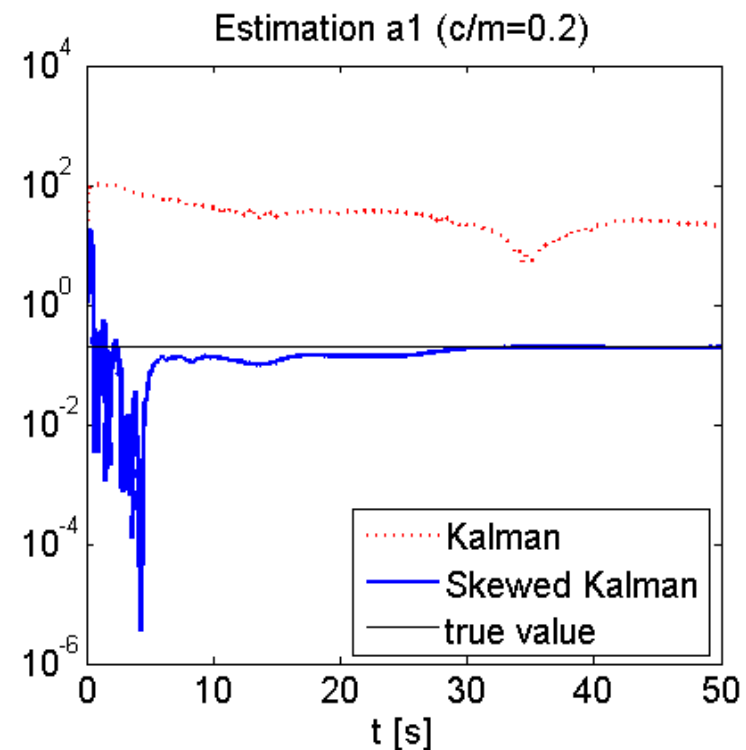
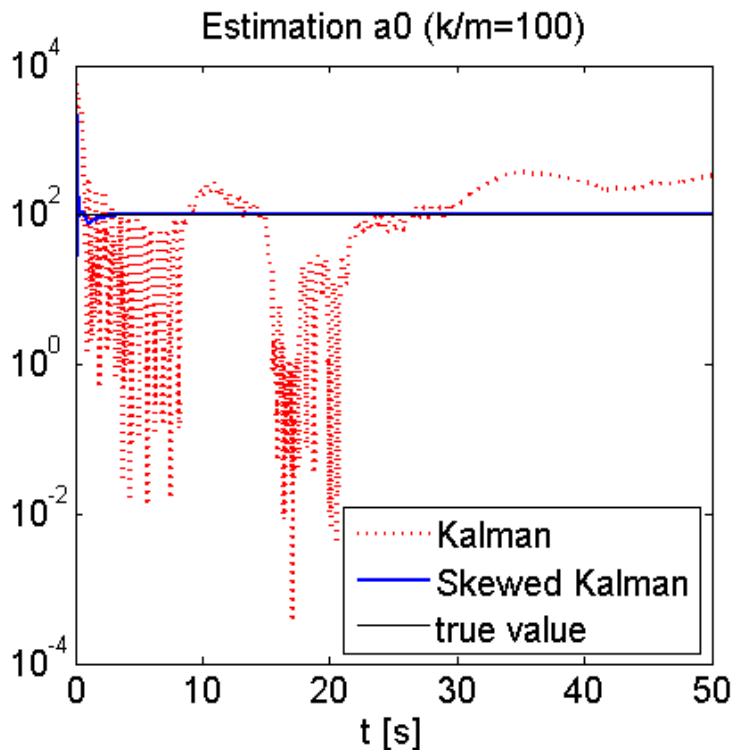


# Approximation of the PDF of the measurement wind data



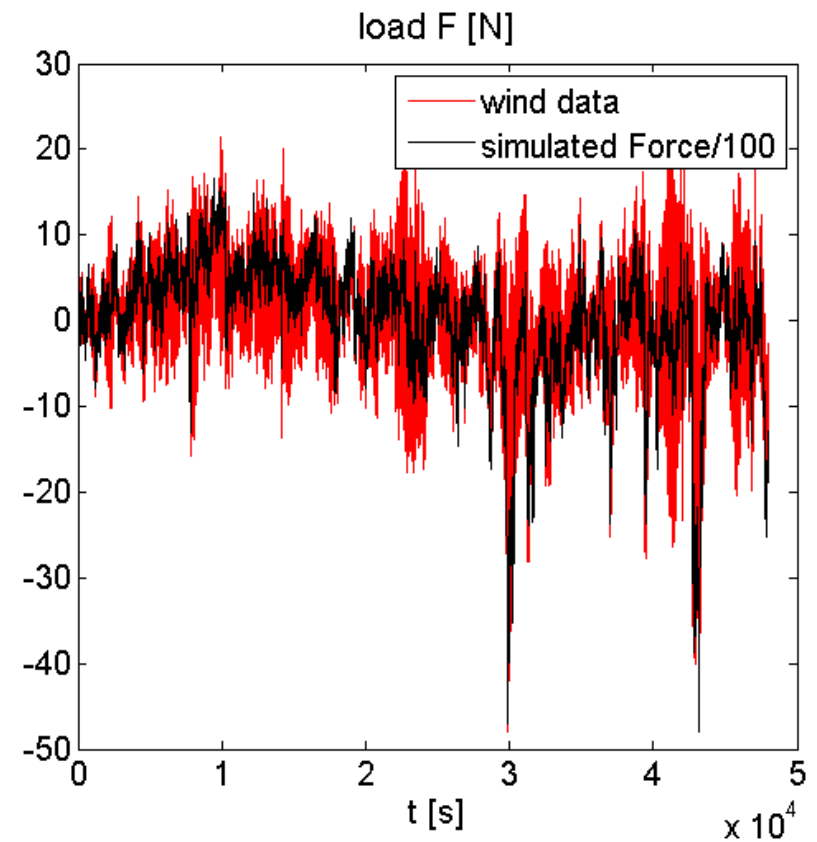
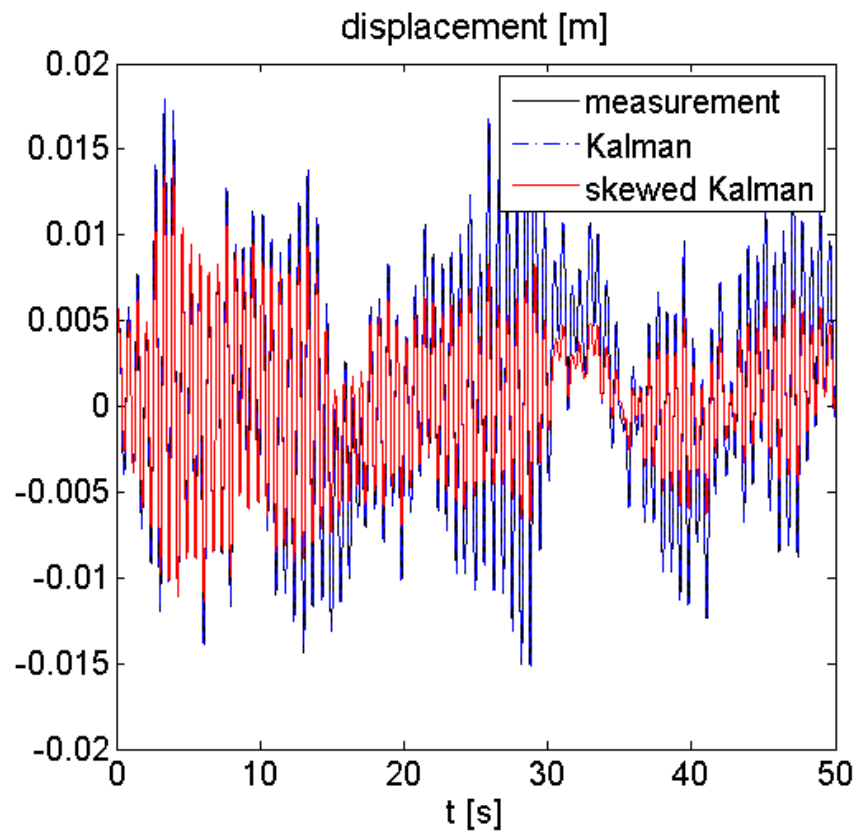
# Identification results

	true values	initial values	initial error	identified parameters		identification error	
				EKF	EsKF	EKF	EsKF
$a_0 [s^{-2}]$	100	150	50 %	20,40	101,84	79,6 %	2 %
$a_1 [s^{-1}]$	0,2	0,3	50 %	1,18	0,19	490 %	5%





# Predicted displacement and generated load



## Conclusion

- ✚ The method is computational efficient
- ✚ The method worked well for the introduced identification problem
- ✚ The CSN allows to describe the characteristics of the regarded skewed wind data
- ✖ The „form“ of the temporal evaluation of the wind data was approximated well, but the amplitude was wrong by a factor 100

## Further work

- Find the error....any comments?
- Investigate the sensitivity of the parameter identification results depending on the initial conditions
- Extending the method to a multi-degree of freedom system

Thanks!  
Any questions or comments?

