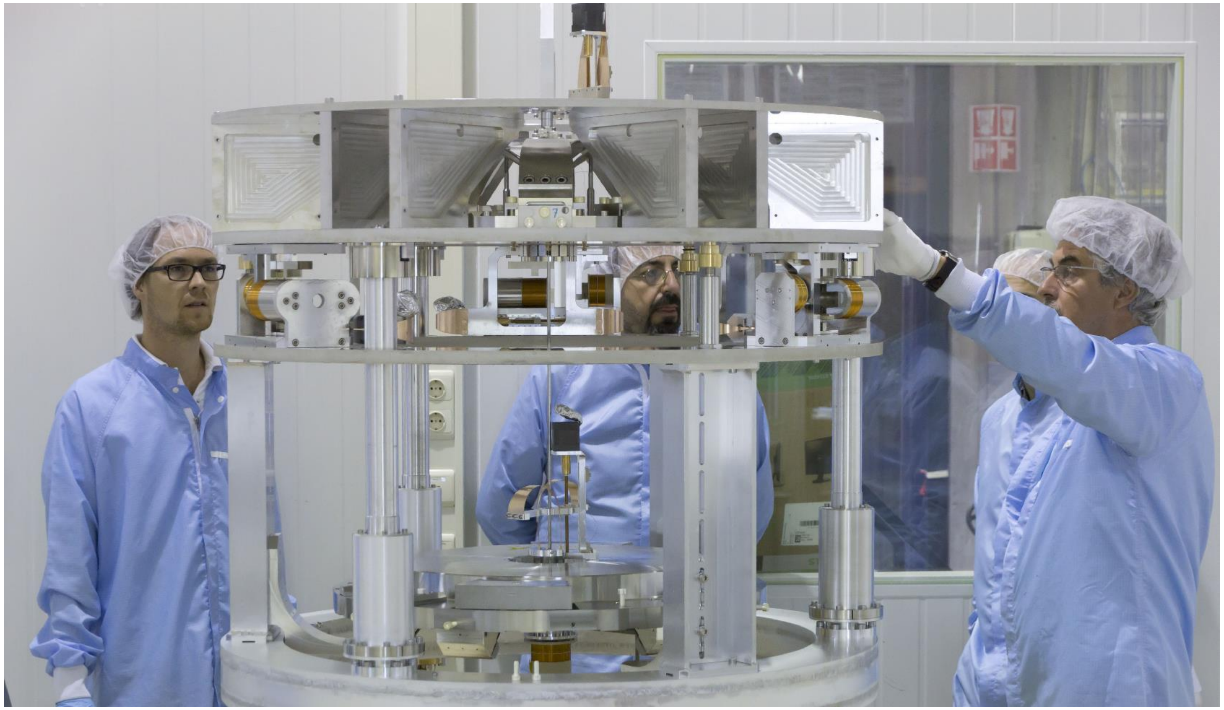
State observers and Kalman filtering

High performance vibration isolation systems

Prof.dr. J.F.J. van den Brand



Nikhef

GWADW, May 19, 2015

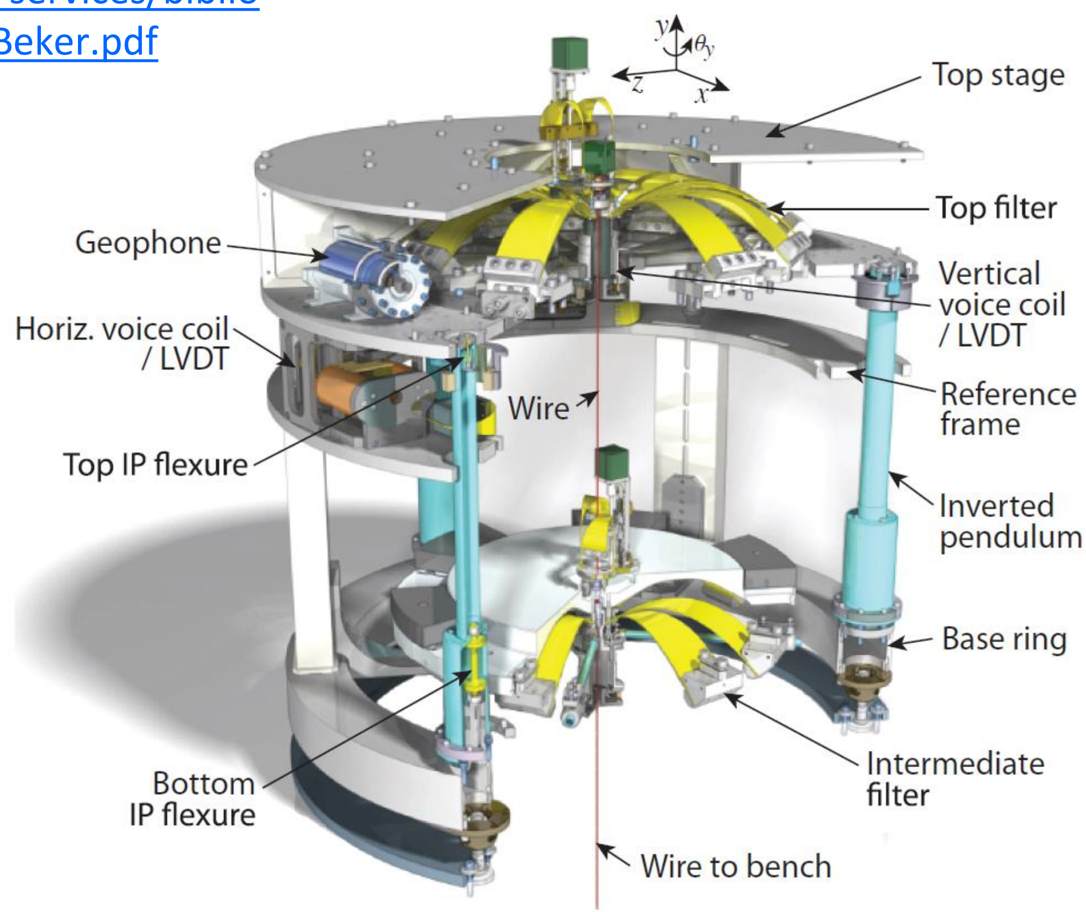
jo@nikhef.nl

Multi-stage attenuation systems

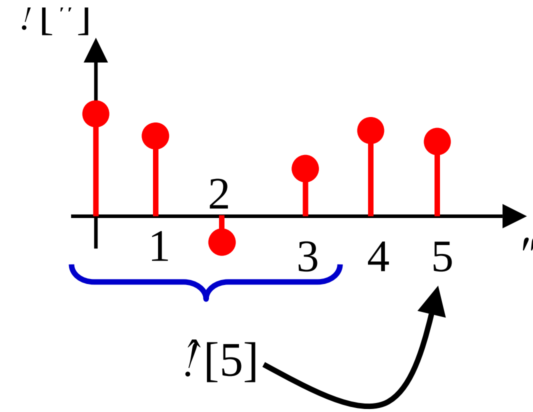
- Bench low frequency control
  - Mark Beker
    - Paper, see:
    - http://www.nikhef.nl/pub/services/biblio/theses_pdf/thesis_M_G_Beker.pdf
  - Vertical DOF

State observers and Kalman filtering for high performance vibration isolation systems

M.G. Beker,[1,*] A. Bertolini,[1] J.F.J. van den Brand,[1,2] H.J. Bulten,[1,2] E. Hennes,[1] and D.S. Rabeling[1]
[1]National Institute for Subatomic Physics Nikhef, Science Park 105, 1098 XG, Amsterdam, The Netherlands
[2]VU University Amsterdam, de Boelelaan 1081, 1081 HV Amsterdam, The Netherlands

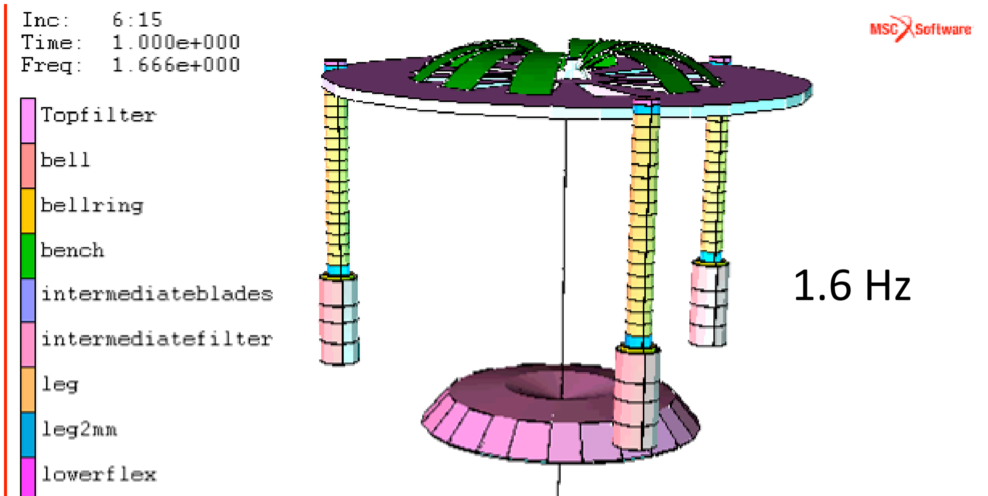


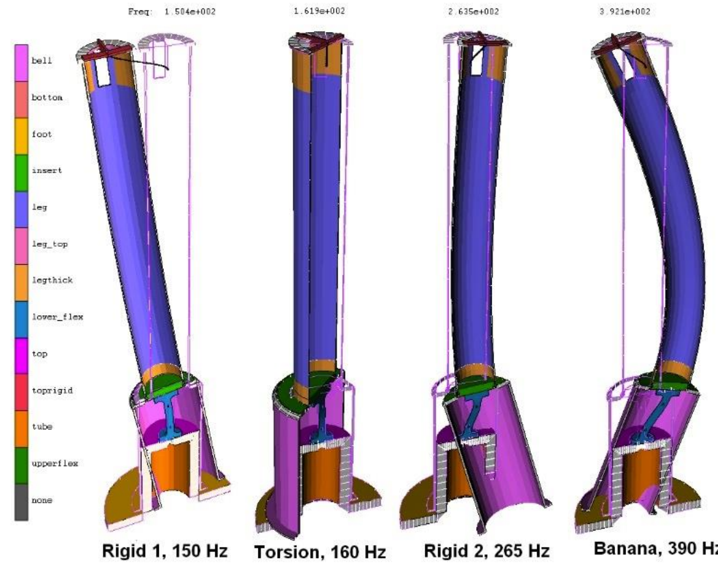
Wiener vs Kalman filtering



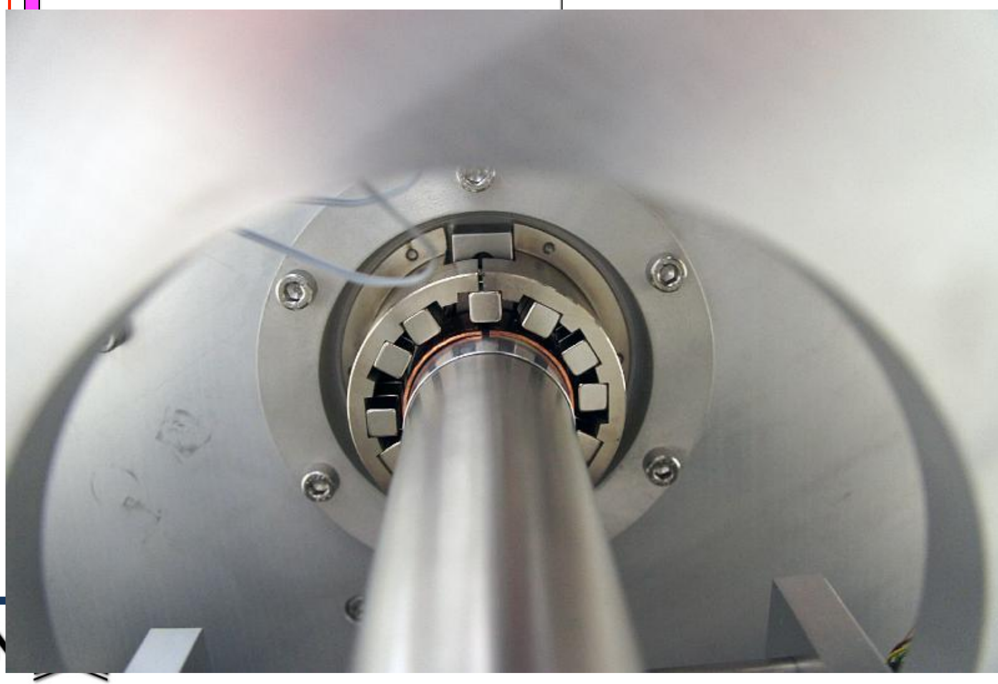
- Least squares
  - Minimizes the sum of squares of the errors
  - Has no “knowledge” of the system
- Wiener filter
  - x[n] = s[n] + w[n] -> “estimate s[n] so as to minimize the error”
  - Stationary processes – The statistical properties of the inputs don’t change in t
  - Causal, length grows, (generally) non-recursive
  - For discrete samples reduces to least squares solution
- Kalman filter
  - Generalization for Wiener filter to non-stationary processes – The signal is characterized by a dynamical model
  - Recursive – don’t need to re-evaluate all data at each step
  - Uses prior knowledge of the system
    - System is described by state vector **x** (unobservable)
    - State can be estimated based on ẍ previous data **z** and model
    - Requires a dynamic (state space) model

Finite element models used to identify all modes of the system



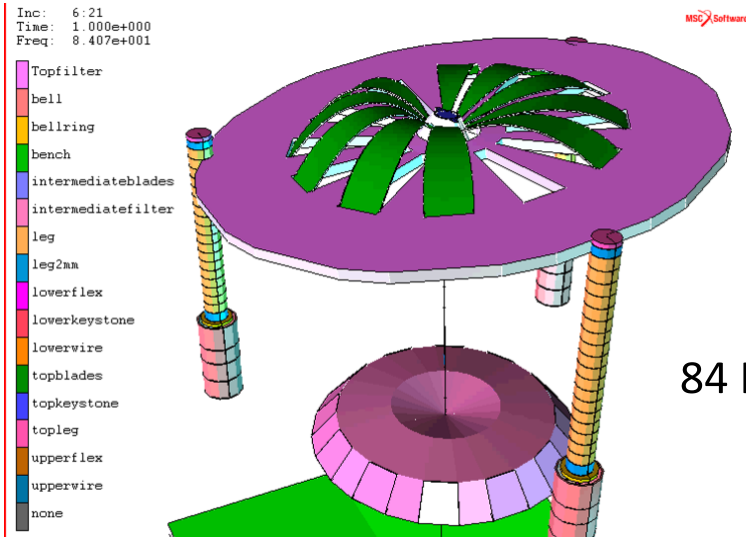


IP leg modes



Where needed passive eddy current dampers can be used to lower Q-factor of higher order resonances

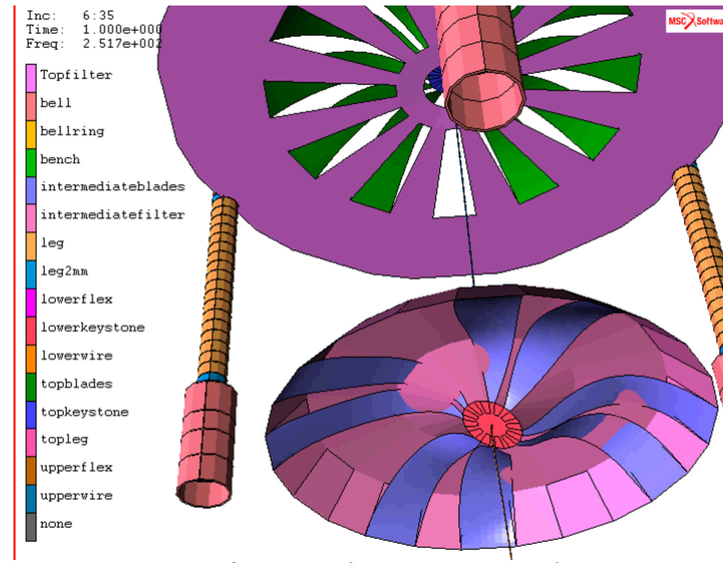
Finite element model: Higher order modes



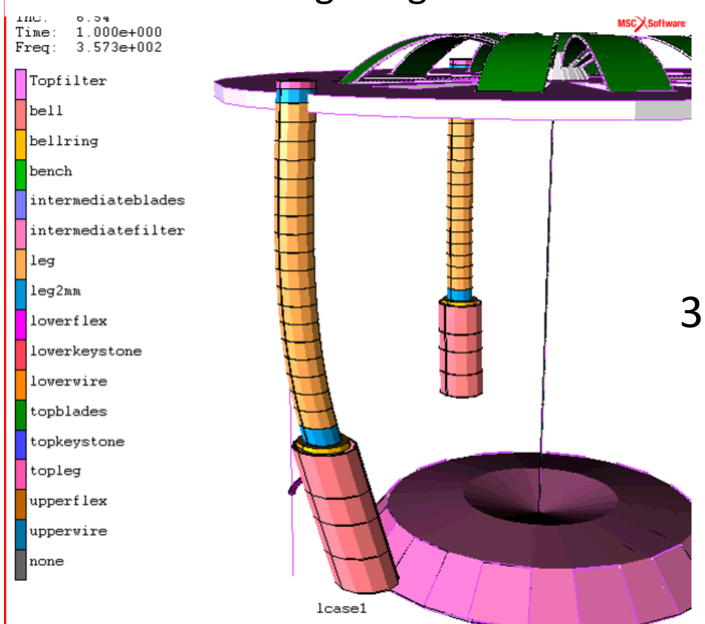
Non-rigid leg modes

84 Hz

252 Hz

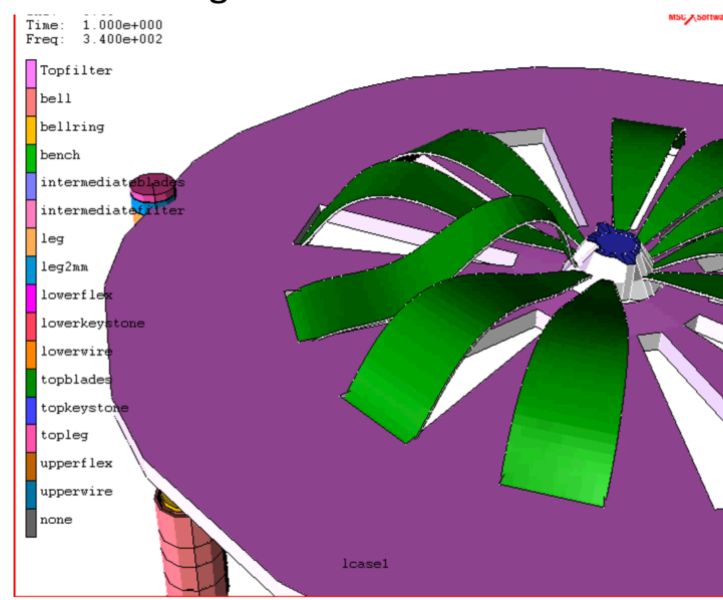


Higher order GAS modes

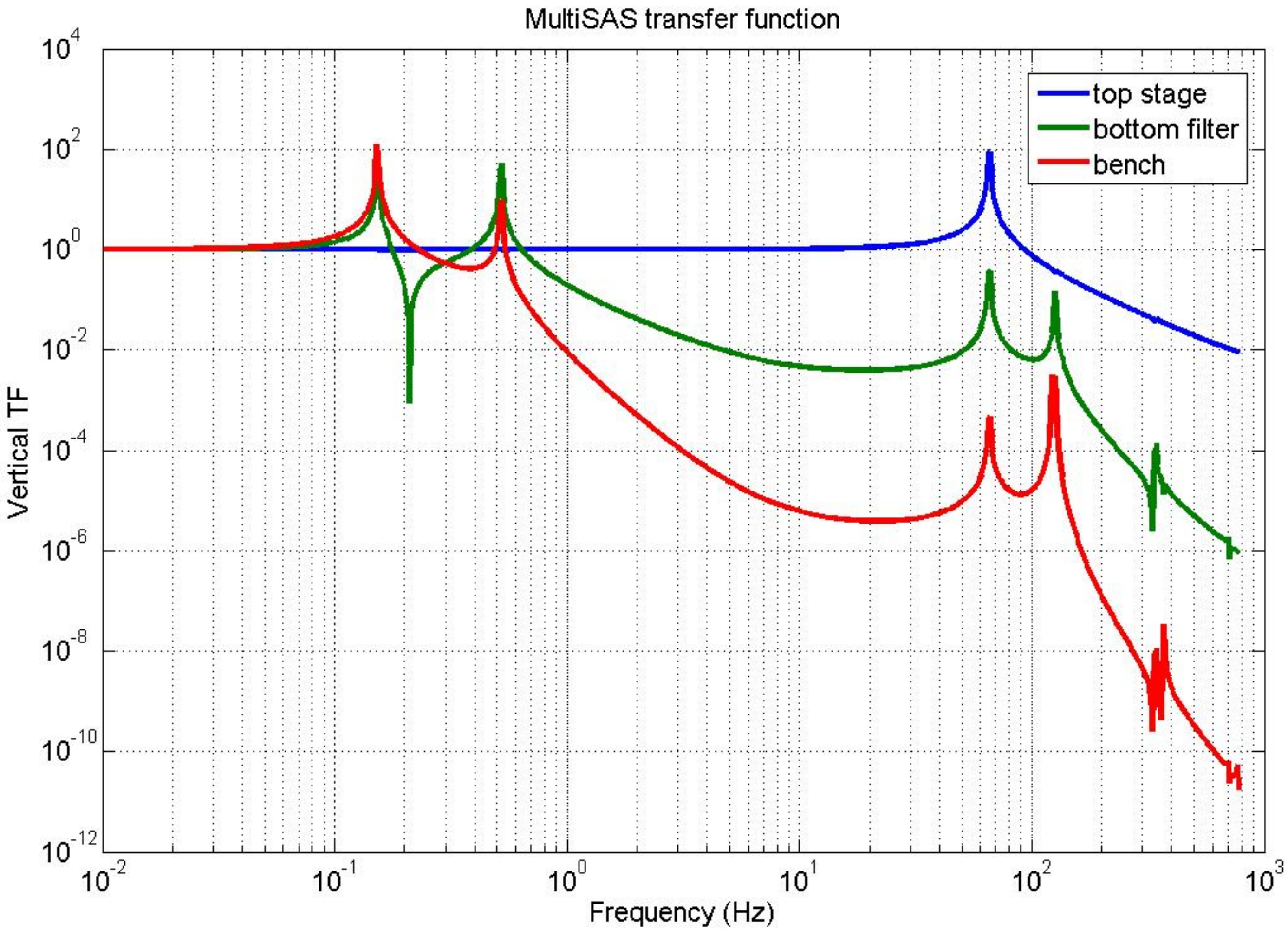


357 Hz

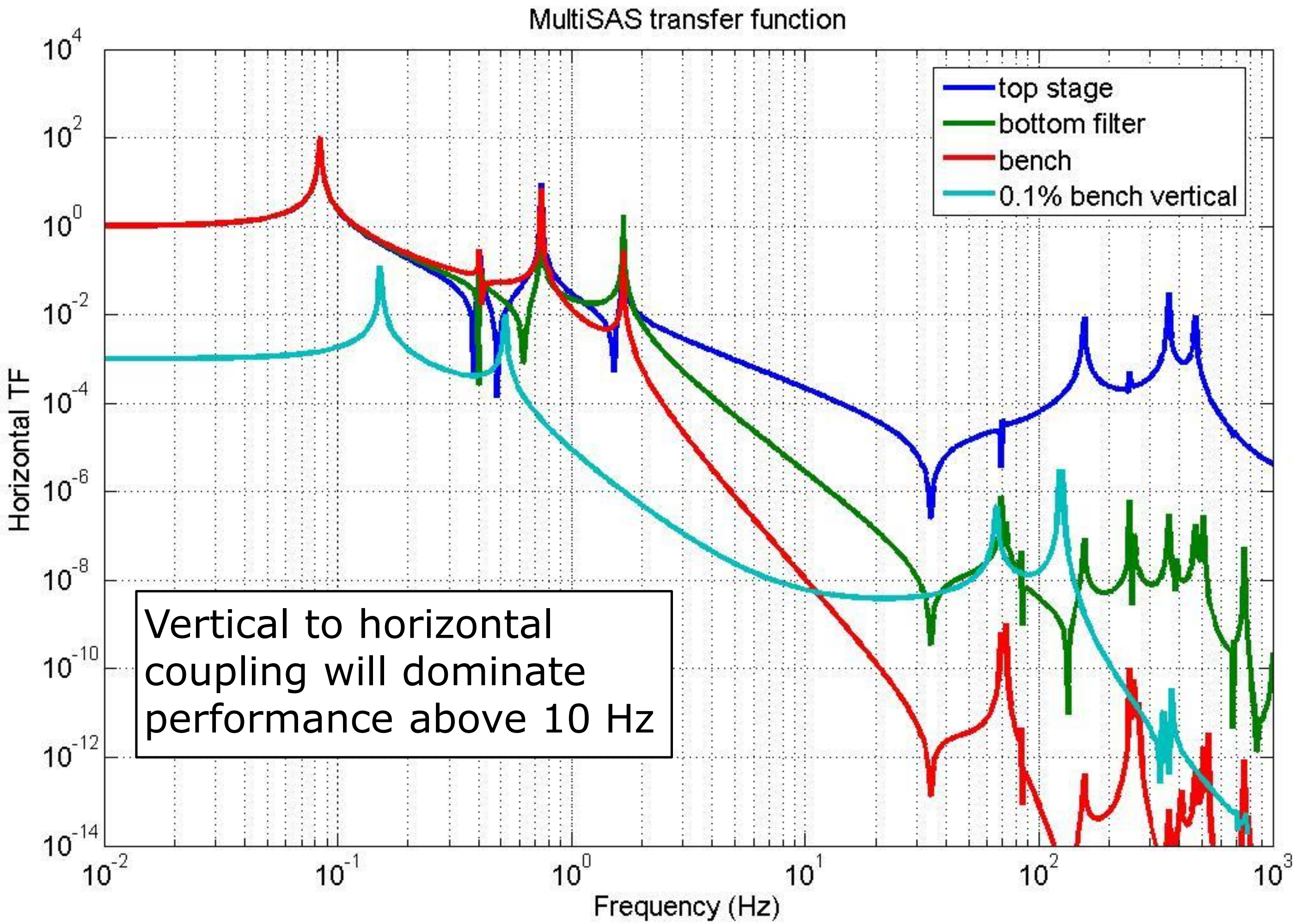
340 Hz



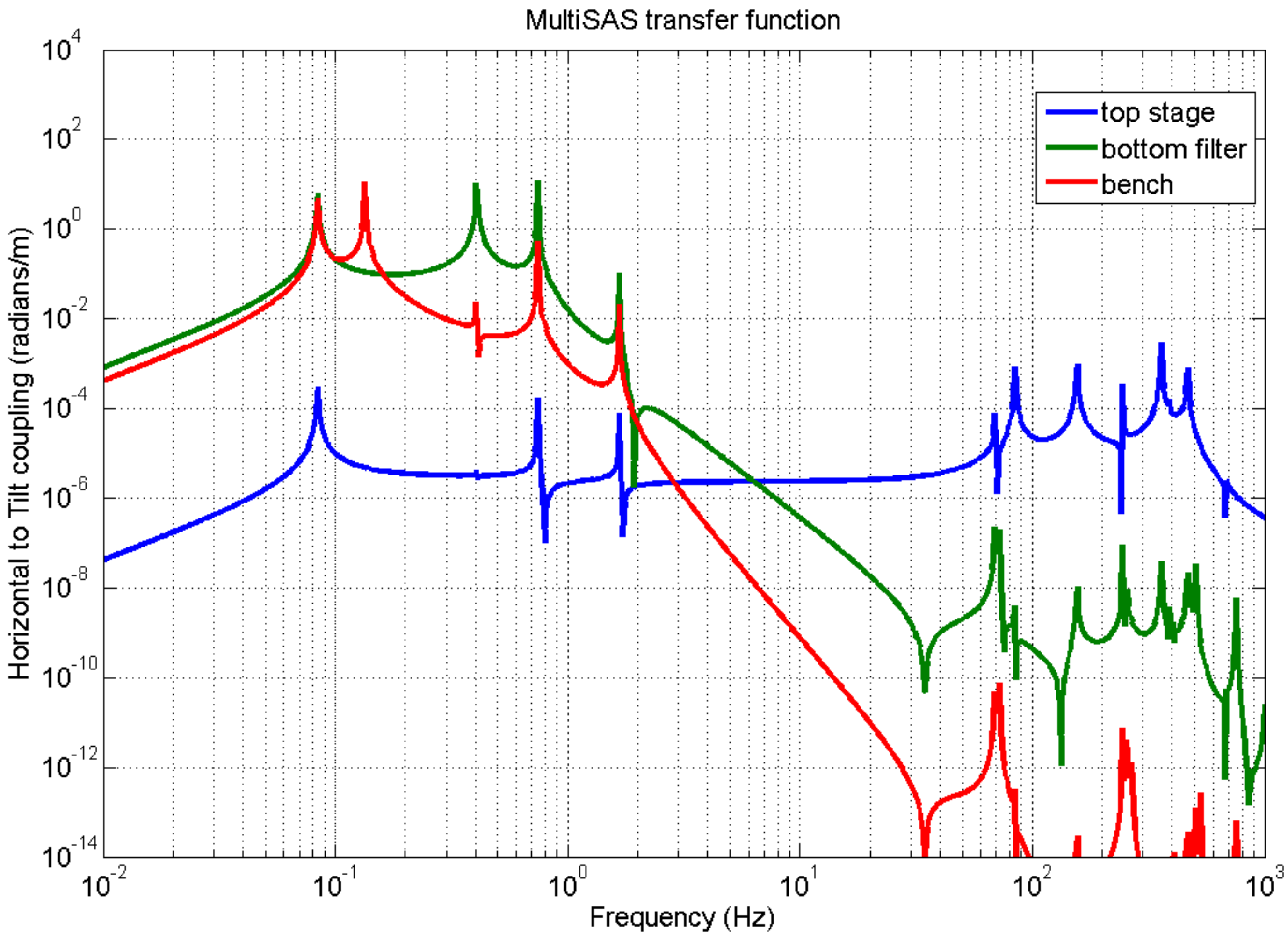
Simulated performance
Vertical transfer function



Simulated performance
Horizontal transfer function



Simulated performance
Horizontal to tilt coupling



Optimal control: state observer

- State space model
  - Imperative to have an accurate model
- FEA
  - Detailed description of the system
  - Tune model to measured transfer functions


Questions?

SYSTEM IDENTIFICATION WITH BAYES THEOREM AND NON GAUSSIAN DISTRIBUTIONS

G. Cella

Three basic rules

1. Law of total probability

$$P\left(\boldsymbol{x}_n, t_n \mid \boldsymbol{y}_1, t_1; \cdots; \boldsymbol{y}_{n-1}, t_{n-1}\right) = \int P\left(\boldsymbol{x}_n, t_n \mid \boldsymbol{x}_{n-1}, t_{n-1}\right) P\left(\boldsymbol{x}_{n-1}, t_{n-1} \mid \boldsymbol{y}_1, t_1; \cdots; \boldsymbol{y}_{n-1}, t_{n-1}\right) d\boldsymbol{x}_{n-1}$$

2. Bayes' theorem

$$P\left(\boldsymbol{x}_n, t_n \mid \boldsymbol{y}_1, t_1; \cdots; \boldsymbol{y}_n, t_n\right) = \frac{P\left(\boldsymbol{y}_n \mid \boldsymbol{x}_n\right) P\left(\boldsymbol{x}_n, t_n \mid \boldsymbol{y}_1, t_1; \cdots; \boldsymbol{y}_{n-1}, t_{n-1}\right)}{\int P\left(\boldsymbol{y}_n \mid \boldsymbol{x}_n\right) P\left(\boldsymbol{x}_n, t_n \mid \boldsymbol{y}_1, t_1; \cdots; \boldsymbol{y}_{n-1}, t_{n-1}\right) d\boldsymbol{x}_n}$$

3. The product of two (multivariate) gaussian distributions is proportional to a (multivariate) gaussian distributions:

$$\mathcal{N}\left(\boldsymbol{\mu}_1, \mathbb{C}_1^{-1}\right) \mathcal{N}\left(\boldsymbol{\mu}_2, \mathbb{C}_2^{-1}\right) = Z_{12} \mathcal{N}\left(\boldsymbol{\mu}_{12}, \mathbb{C}_{12}^{-1}\right)$$

$$\mathbb{C}_{12}^{-1} = \mathbb{C}_1^{-1} + \mathbb{C}_2^{-1}$$

$$\boldsymbol{\mu}_{12} = \mathbb{C}_{12}\left(\mathbb{C}_1^{-1}\boldsymbol{\mu}_1 + \mathbb{C}_2^{-1}\boldsymbol{\mu}_2\right)$$

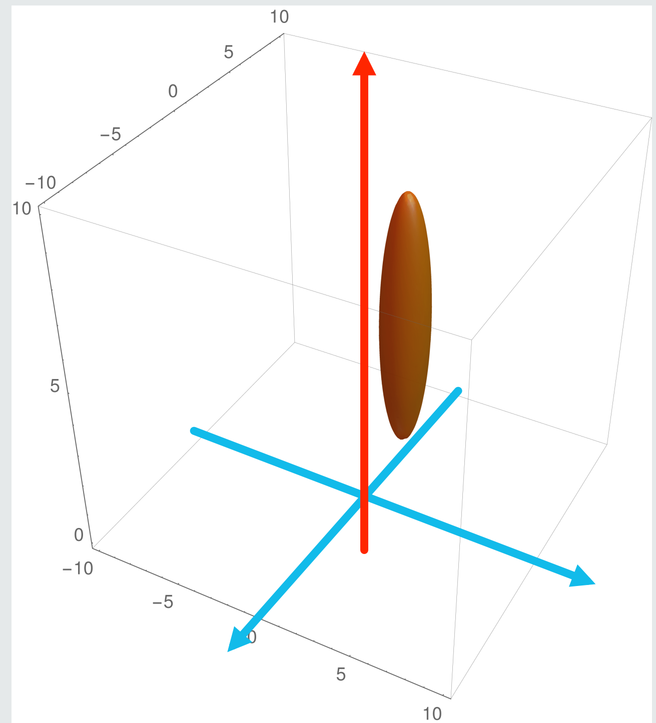
A simple example: a suspension (pendulum) with uncertain length

- Description in the phase space:

$$\dot{p} = -m\omega^2 x$$

$$\dot{x} = \frac{1}{m} p$$

- We measure the state (position and velocity). Maybe with some measurement error.
- We enlarge the space, adding the unknown parameter ω
- We model our ignorance with a joint probability distribution
- We assume we have a good model…
- …which can be used to calculate the time evolution (**RULE 1 at work**)



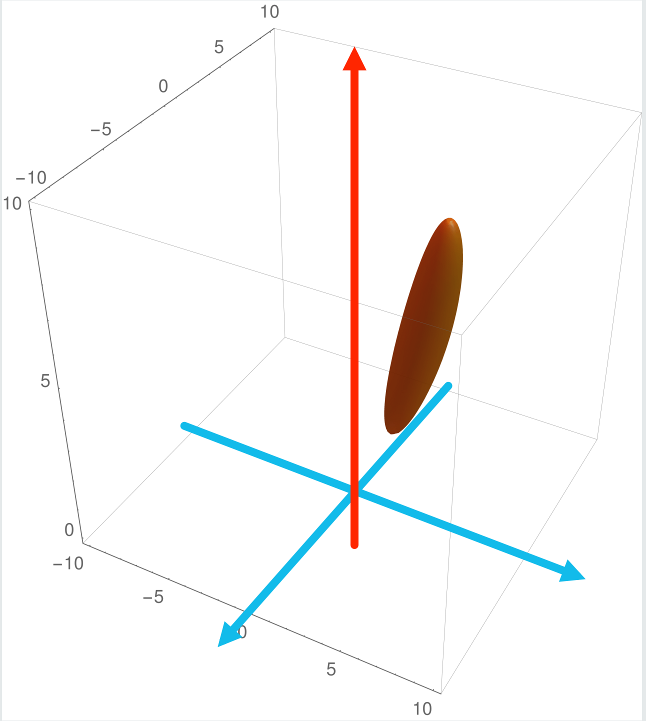
A simple example: a suspension (pendulum) with uncertain length

- Description in the phase space:

$$\dot{p} = -m\omega^2 x$$

$$\dot{x} = \frac{1}{m} p$$

- We measure the state (position and velocity). Maybe with some measurement error.
- We enlarge the space, adding the unknown parameter ω
- We model our ignorance with a joint probability distribution
- We assume we have a good model…
- …which can be used to calculate the time evolution (**RULE 1 at work**)



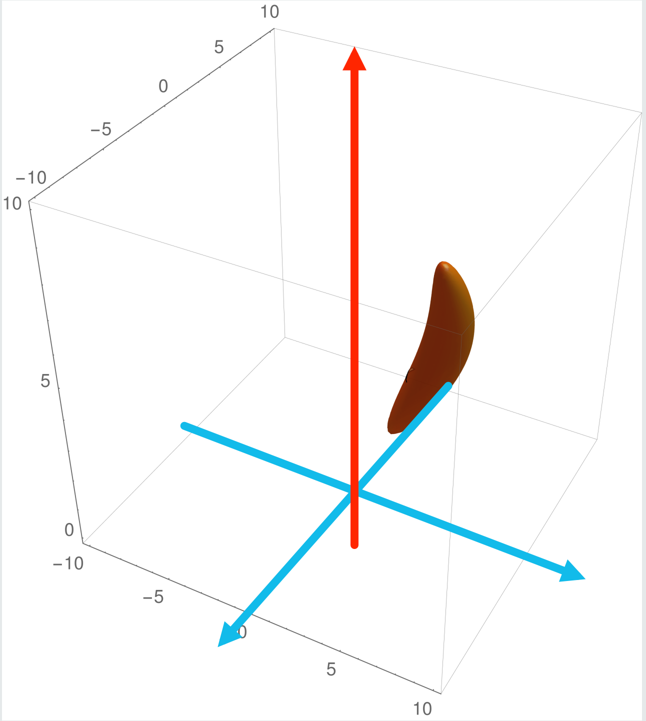
A simple example: a suspension (pendulum) with uncertain length

- Description in the phase space:

$$\dot{p} = -m\omega^2 x$$

$$\dot{x} = \frac{1}{m} p$$

- We measure the state (position and velocity). Maybe with some measurement error.
- We enlarge the space, adding the unknown parameter ω
- We model our ignorance with a joint probability distribution
- We assume we have a good model…
- …which can be used to calculate the time evolution (**RULE 1 at work**)



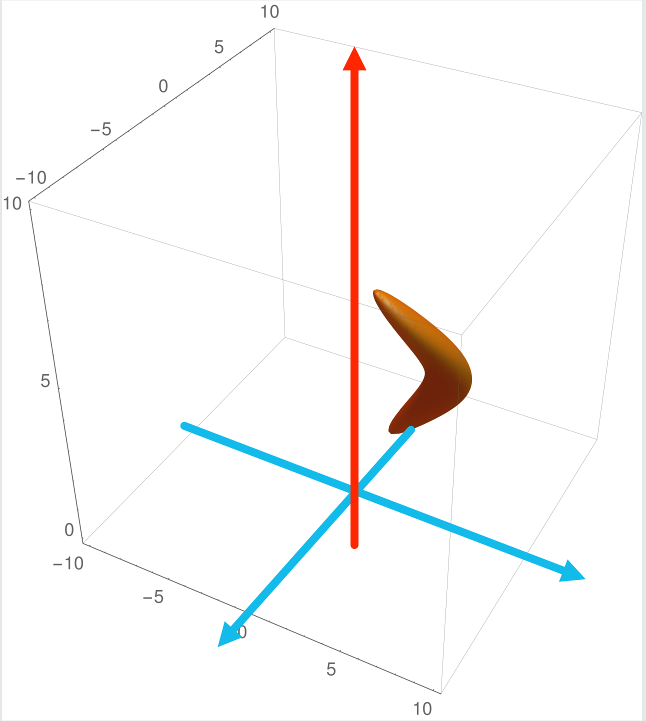
A simple example: a suspension (pendulum) with uncertain length

- Description in the phase space:

$$\dot{p} = -m\omega^2 x$$

$$\dot{x} = \frac{1}{m} p$$

- We measure the state (position and velocity). Maybe with some measurement error.
- We enlarge the space, adding the unknown parameter ω
- We model our ignorance with a joint probability distribution
- We assume we have a good model…
- …which can be used to calculate the time evolution (**RULE 1 at work**)



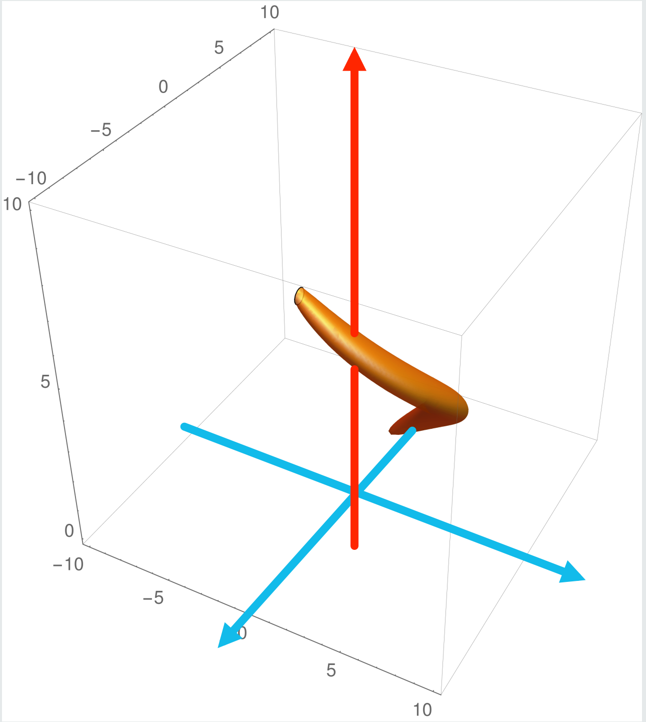
A simple example: a suspension (pendulum) with uncertain length

- Description in the phase space:

$$\dot{p} = -m\omega^2 x$$

$$\dot{x} = \frac{1}{m} p$$

- We measure the state (position and velocity). Maybe with some measurement error.
- We enlarge the space, adding the unknown parameter ω
- We model our ignorance with a joint probability distribution
- We assume we have a good model…
- …which can be used to calculate the time evolution (**RULE 1 at work**)



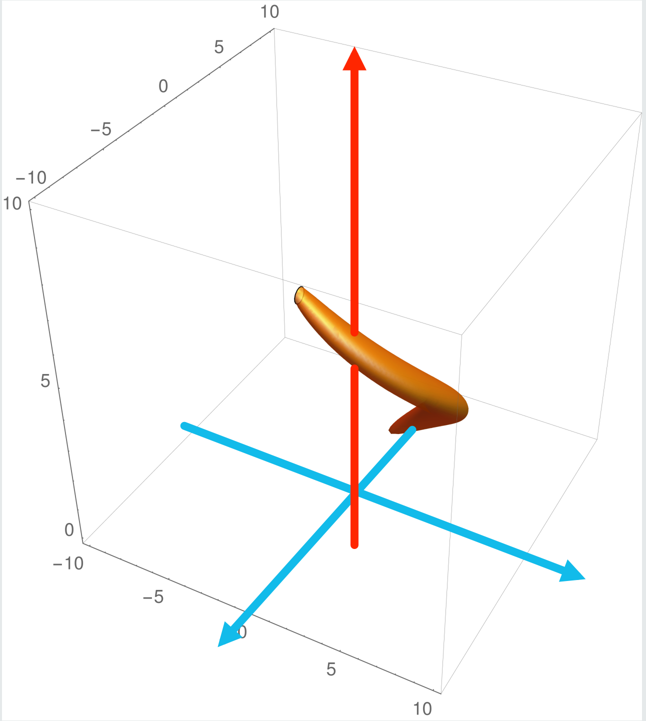
A simple example: a suspension (pendulum) with uncertain length

- This is no more a gaussian distribution (in general). How to parameterize it?
  - Each horizontal line is a gaussian distribution
  - Gaussian misture can be a good representation:

$$\sim \sum w_i \mathcal{N}(\boldsymbol{\mu}_i, \mathbb{C}_i)$$

- Now, we measure the position and the velocity again, and we use **RULE 2** and **RULE 3**

$$\sim \sum w_i^* \mathcal{N}(\boldsymbol{\mu}_i^*, \mathbb{C}_i^*)$$



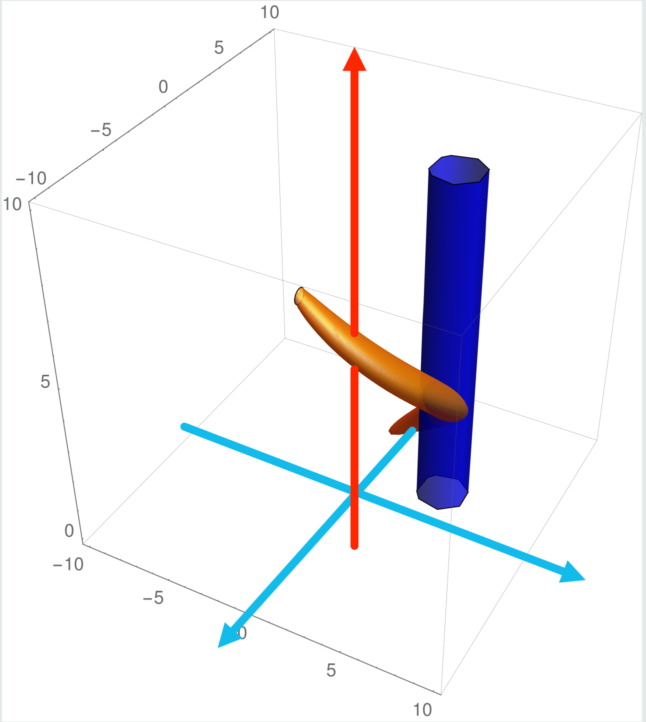
A simple example: a suspension (pendulum) with uncertain length

- This is no more a gaussian distribution (in general). How to parameterize it?
  - Each horizontal line is a gaussian distribution
  - Gaussian misture can be a good representation:

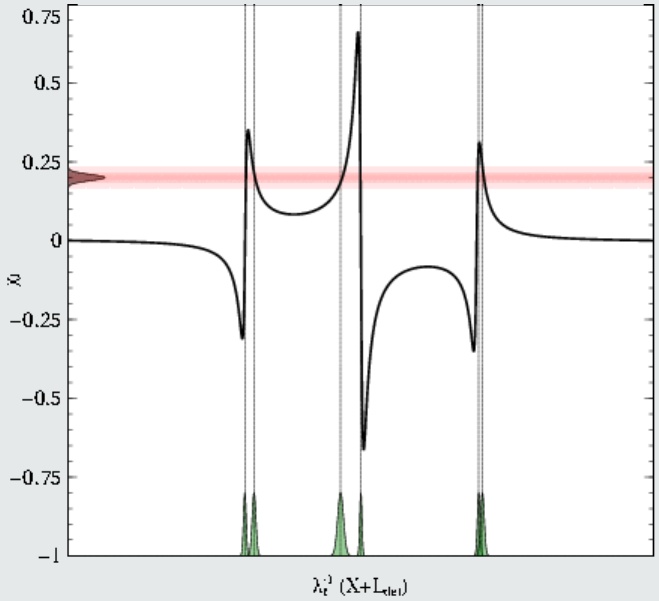
$$\sim \sum w_i \mathcal{N}(\boldsymbol{\mu}_i, \mathbb{C}_i)$$

- Now, we measure the position and the velocity again, and we use **RULE 2** and **RULE 3**

$$\sim \sum w_i^* \mathcal{N}(\boldsymbol{\mu}_i^*, \mathbb{C}_i^*)$$



This could have several applications



- Tracking the Pound Drever signal
  - Identify optical parameters
  - Improve locking
- Systems with nonlinear dynamics
  - Radiation pressure
- Adjusting noise models:

$$dZ(t) = \int K(t, t')\, dW(t')$$

- Selection strategy needed
  - Elements with low weight must be removed
  - Gaussian misture not necessarily a good representation in all cases

Suspension Parameter Estimation for State Space Models

Brett Shapiro

GWADW – 19 May 2015

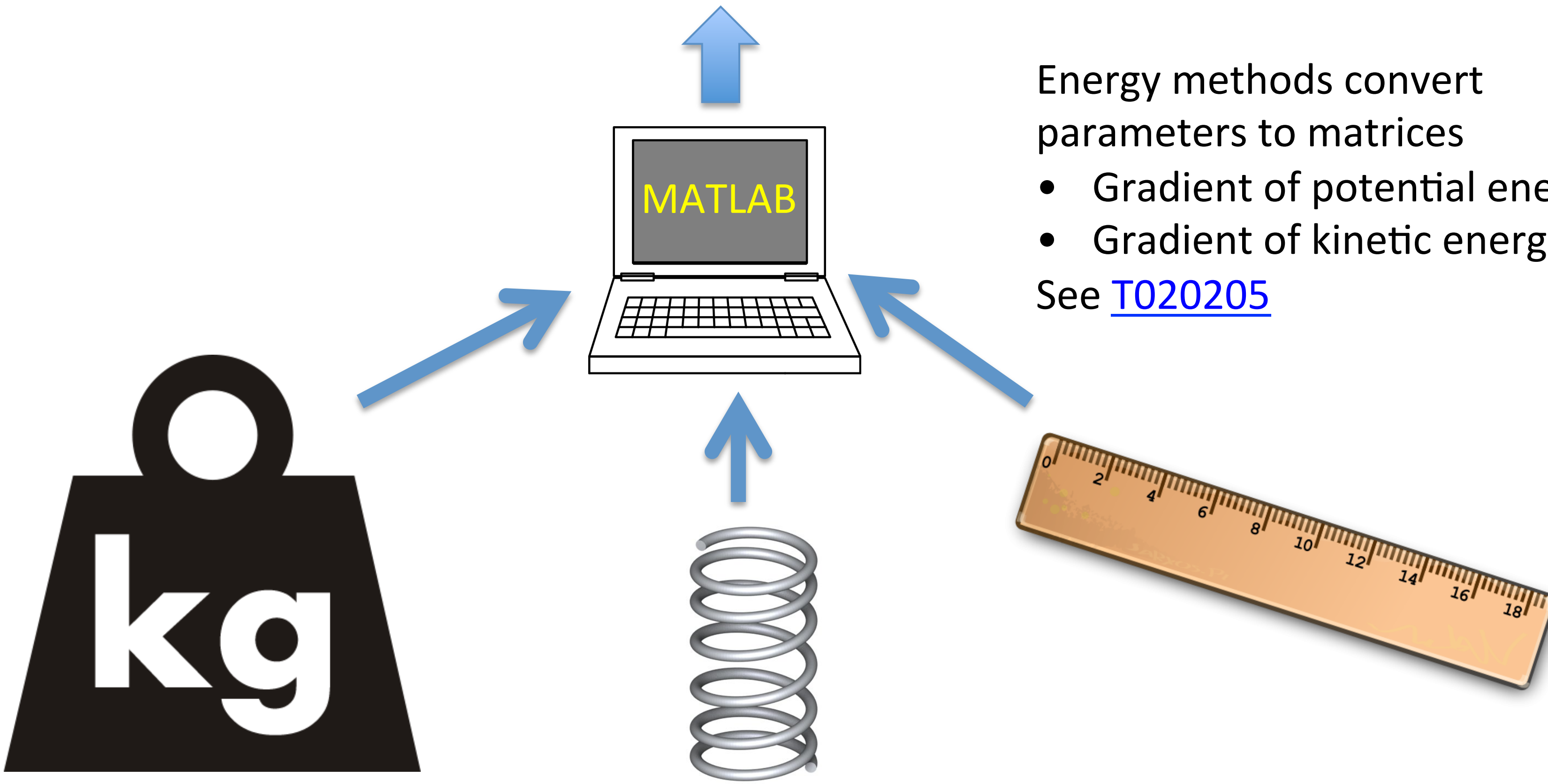
State space from physical values

$$\dot{x} = Ax + Bu$$

Energy methods convert parameters to matrices

- Gradient of potential energy
- Gradient of kinetic energy

See [T020205]



Inertia

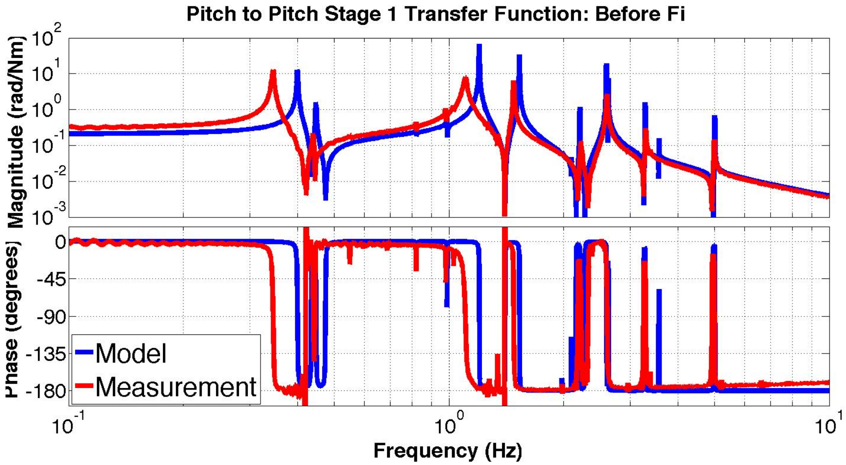
Stiffness

Length

Parameter Estimation Algorithm

$$J = \begin{bmatrix} \frac{\partial e_1}{\partial p_1} & \frac{\partial e_1}{\partial p_2} & \cdots \\ \frac{\partial e_2}{\partial p_1} & \frac{\partial e_2}{\partial p_2} & \cdots \\ \frac{\partial e_3}{\partial p_1} & \frac{\partial e_3}{\partial p_2} & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}$$

- J = Jacobian matrix of error gradients wrt parameters
- e = % error between modeled and measured resonant frequencies
- p = parameter value (mass, stiffness, length, etc)



Mismatch between model and measurement

Parameter Estimation Algorithm

$$J = \begin{bmatrix} \frac{\partial e_1}{\partial p_1} & \frac{\partial e_1}{\partial p_2} & \cdots \\ \frac{\partial e_2}{\partial p_1} & \frac{\partial e_2}{\partial p_2} & \cdots \\ \frac{\partial e_3}{\partial p_1} & \frac{\partial e_3}{\partial p_2} & \cdots \\ \cdots & \cdots & \cdots \end{bmatrix}$$

- *J* = Jacobian matrix of error gradients wrt parameters
- *e* = % error between modeled and measured resonant frequencies
- *p* = parameter value (mass, stiffness, length, etc)

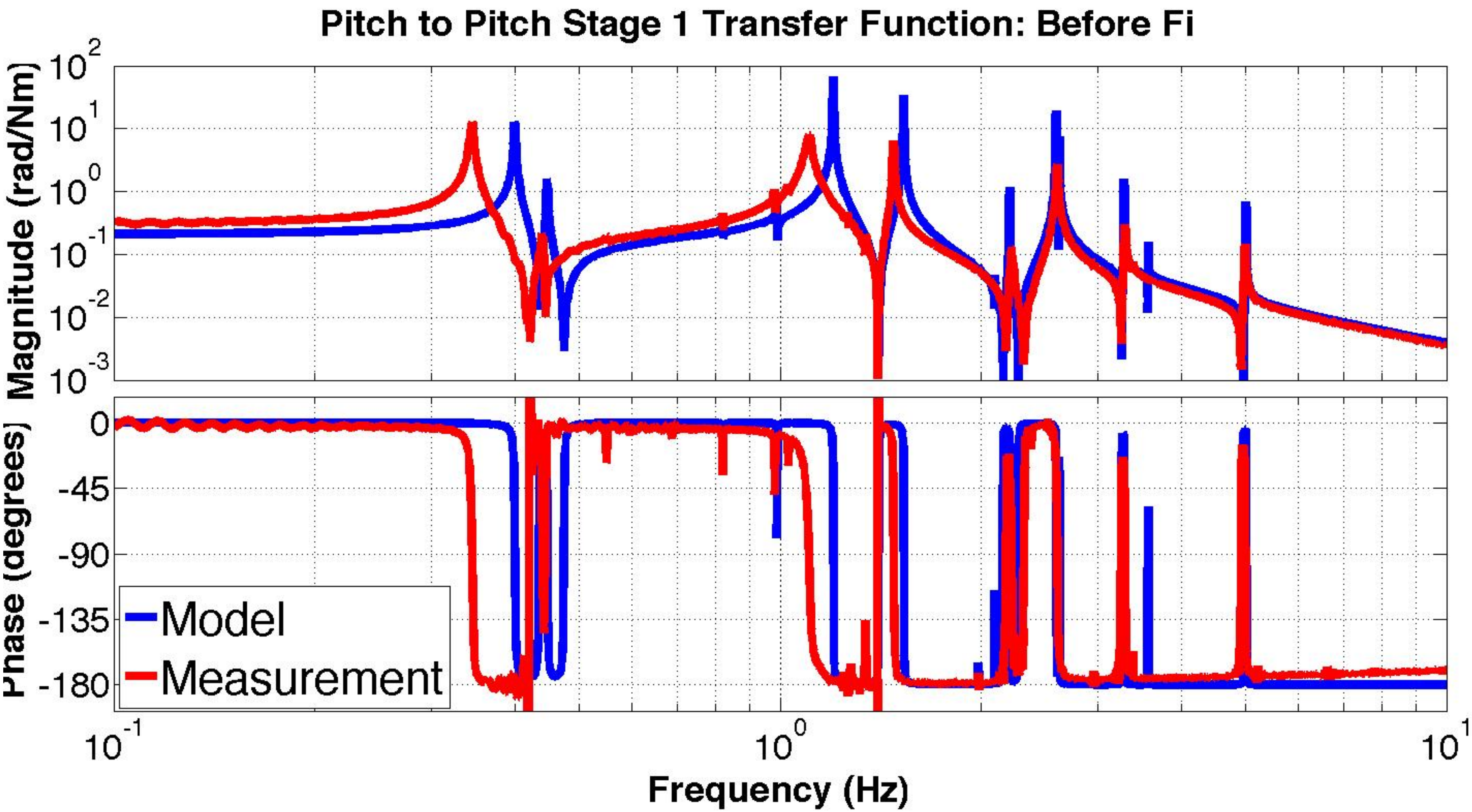
Gauss-Newton algorithm – an modification of Newton's method (2nd order)

$$p_{k+1} = p_k + \alpha_k d_k$$ Update the parameter list, *p*, with step size *α*

$$d_k = J_k^{\dagger} e_k$$ Update the descent direction *d* with the psuedo-inverse of *J*.

References: T1000458 and "Selection of Important Parameters Using Uncertainty and Sensitivity Analysis", Shapiro et al.

Before Parameter Estimation



After Parameter Estimation