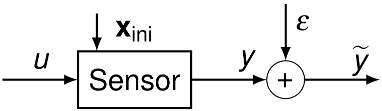
Statistical analysis and experimental validation of data-driven dynamic measurement methods

Gustavo Quintana Carapia

VUB FACULTY OF ENGINEERING

A measurement is a dynamic process that estimates the input from the sensor response

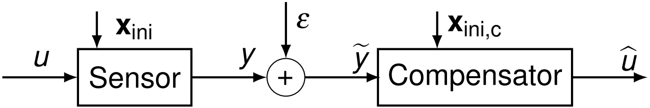
- The sensor is a dynamic system.
- The sensor response is perturbed with errors.



- The aim is to assess the uncertainty of the data-driven input estimation method.

How is the input estimated from the sensor response?

- With sensor model
  - Kalman filter.
  - Compensator system.

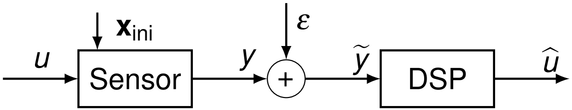


- Without sensor model
  - System identification and input estimation.
  - Optimization.
  - Data-driven methods.

Digital Signal Processors offer an alternative to classical approaches

A DSP, with appropriate algorithms, can

- emulate dynamical systems, or
- implement data-driven methods.



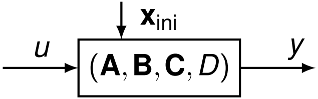
- The statistical properties of data-driven methods are not straightforward evident.

Statistical analysis and experimental validation of data-driven dynamic measurement methods

- Formulation of a data-driven step input estimation method.

- Thesis contribution.
  - Statistical analysis, structured EIV problems.
  - Experimental validation.
  - Affine input estimation methods.

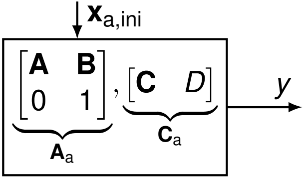
The input estimation methods are formulated as output-error problems



- With model and exact data, we can write

$$\underbrace{\begin{bmatrix} y(0) \\ y(1) \\ y(2) \\ \vdots \\ y(N) \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^N \end{bmatrix}}_{\mathcal{O}} \mathbf{x}(0) + \underbrace{\begin{bmatrix} \mathbf{D} & & & & \\ \mathbf{CB} & \mathbf{D} & & & \\ \mathbf{CAB} & \mathbf{CB} & \mathbf{D} & & \\ \vdots & & \ddots & & \\ \mathbf{CA}^{N-1}\mathbf{B} & \cdots & \mathbf{CAB} & \mathbf{CB} & \mathbf{D} \end{bmatrix}}_{\mathcal{T}} \underbrace{\begin{bmatrix} u(0) \\ u(1) \\ u(2) \\ \vdots \\ u(N) \end{bmatrix}}_{\mathbf{u}}$$

With a step input, the sensor admits an augmented autonomous state space representation



- Without sensor model, we can estimate first $\mathcal{O}_{\mathrm{a}}$ from

$$\underbrace{\begin{bmatrix} y(1) & y(2) & \cdots & y(n) \\ y(2) & y(3) & \cdots & y(n+1) \\ \iddots & \iddots & \iddots & \\ y(n) & y(n+1) & \cdots & y(2n-1) \end{bmatrix}}_{\mathcal{H}(y)} = \underbrace{\begin{bmatrix} \mathbf{C}_{\mathrm{a}} \\ \mathbf{C}_{\mathrm{a}}\mathbf{A}_{\mathrm{a}} \\ \vdots \\ \mathbf{C}_{\mathrm{a}}\mathbf{A}_{\mathrm{a}}^{n} \end{bmatrix}}_{\mathcal{O}_{\mathrm{a}}} \underbrace{\begin{bmatrix} \mathbf{x}_{\mathrm{a}}(1) & \mathbf{x}_{\mathrm{a}}(2) & \cdots & \mathbf{x}_{\mathrm{a}}(n) \end{bmatrix}}_{\mathbf{X}_{\mathrm{ini}}}$$

After estimating $\widehat{\mathcal{O}}_\mathrm{a}$, we can write the total response of the system

$$y = G\,u + \widehat{\mathcal{O}}_\mathrm{a}\,\mathbf{x}_\mathrm{a}(0),$$

- that is equivalent to

$$\underbrace{\begin{bmatrix} y(0) \\ \vdots \\ y(N) \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} \mathbf{1}_{N+1} \otimes G & \widehat{\mathcal{O}}_\mathrm{a} \end{bmatrix}}_{\mathbf{K}} \begin{bmatrix} u \\ \mathbf{x}_\mathrm{a}(0) \end{bmatrix}$$

- where G is the sensor static gain.

To estimate directly the step input, obtain the first difference of the state-space representation

$\downarrow \; \Delta\mathbf{x}_{\text{ini}} = \Delta\mathbf{x}(0)$

$$\boxed{\begin{aligned} \Delta\mathbf{x}(k+1) &= \mathbf{A}\Delta\mathbf{x}(k) \\ \Delta y(k) &= \mathbf{C}\Delta\mathbf{x}(k) \end{aligned}} \xrightarrow{\;\Delta y\;}$$

where: $\Delta\mathbf{x}(k) = \mathbf{x}(k+1) - \mathbf{x}(k)$

- If Δy is persistently exciting enough, the total response is

$$\underbrace{\begin{bmatrix} y(n+1) \\ \vdots \\ y(N) \end{bmatrix}}_{\mathbf{y}} = \underbrace{\begin{bmatrix} \mathbf{1}_{N-n} \otimes G & \mathcal{H}(\Delta y) \end{bmatrix}}_{\mathbf{K}} \underbrace{\begin{bmatrix} u \\ \ell \end{bmatrix}}_{\boldsymbol{\theta}}$$

The data-driven step input estimation method converts the output-error into an errors-in-variables problem

$$\underbrace{\begin{bmatrix} \tilde{y}(n+1) \\ \vdots \\ \tilde{y}(N) \end{bmatrix}}_{\tilde{\mathbf{y}}} = \underbrace{\begin{bmatrix} \mathbf{1}_{N-n} \otimes G & \mathcal{H}\left(\Delta \tilde{y}\right) \end{bmatrix}}_{\tilde{\mathbf{K}}} \underbrace{\begin{bmatrix} u \\ \ell \end{bmatrix}}_{\boldsymbol{\theta}}$$

- additive measurement noise $\tilde{\mathbf{y}} = \mathbf{y} + \epsilon$, and $\tilde{\mathbf{K}} = \mathbf{K} + \mathbf{E}$,
- structured and correlated,
- no evident statistical properties,
- hard to get unbiased solution,
- admits a least-squares solution.

The data-driven step input estimation uncertainty assessment was pending

This requires to find

- the structured errors-in-variables problem
  - stochastic properties,
- the data-driven step input estimation
  - statistical moments,
  - mean-squared-error,

- The uncertainty assessment
  - provides confidence bounds w.r.t sample size,
  - fosters the method utilization in metrology.

Statistical analysis and experimental validation of data-driven dynamic measurement methods

- Formulation of a data-driven step input estimation method.

- Thesis contribution.
  - Statistical analysis, structured EIV problems.
  - Experimental validation.
  - Affine input estimation methods.

To study the least-squares solution of an errors-in-variables problem

$$\widehat{\boldsymbol{\theta}} = (\widetilde{\mathbf{K}}^\top \widetilde{\mathbf{K}})^{-1} \widetilde{\mathbf{K}}^\top \widetilde{\mathbf{y}},$$

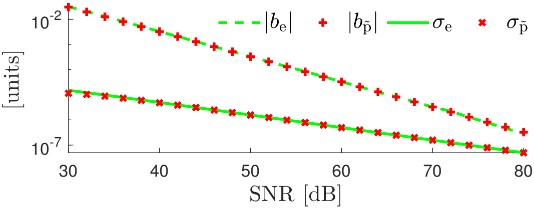
where the noise is assumed additive

$$\widehat{\boldsymbol{\theta}} = \Big((\mathbf{K}+\mathbf{E})^\top (\mathbf{K}+\mathbf{E}) \Big)^{-1} (\mathbf{K}+\mathbf{E})^\top (\mathbf{y}+\boldsymbol{\epsilon}),$$

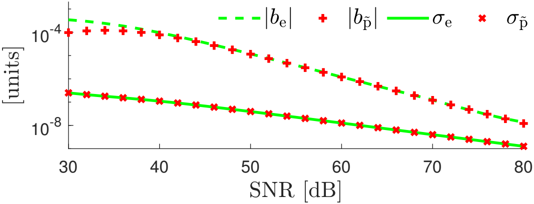
- I calculated
  - the second order Taylor series expansion of $\widehat{\boldsymbol{\theta}}$.
  - expressions that predict the bias and covariance,
    - for unstructured and structured EIV problems.

Monte Carlo simulation shows that the expressions for structured EIV are more susceptible to perturbations

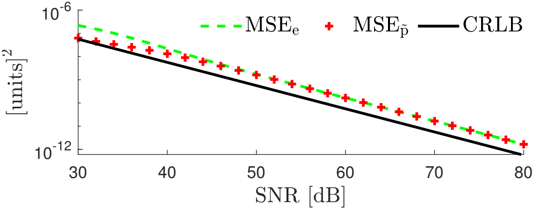
- Unstructured EIV
  - uncorrelated,
  - $\begin{bmatrix}\mathbf{K} & \mathbf{y}\end{bmatrix}$ randomly generated.



- Structured EIV
  - step input estimation from 2nd order model response.



For a 2nd order system, the step input estimation MSE is larger than the CRLB by one order of magnitude

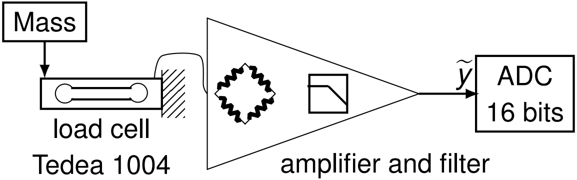


Statistical analysis and experimental validation of data-driven dynamic measurement methods

- Formulation of a data-driven step input estimation method.

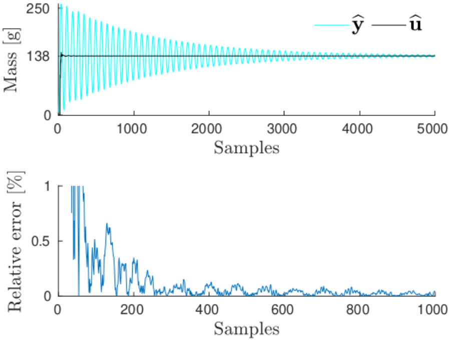
- Thesis contribution.
  - Bias and variance analysis, structured EIV problems.
  - Experimental validation.
  - Affine input estimation methods.

I designed and built a weighing system to collect experimental data from step inputs

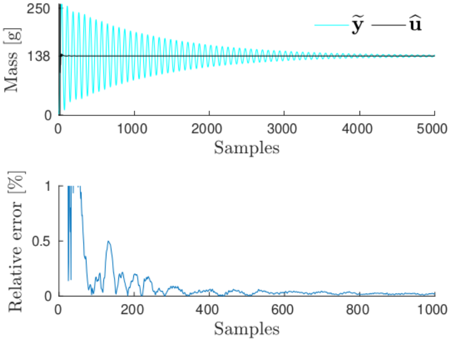


- A 5th order sensor model was obtained with the sysId toolbox.

These are typical step input estimation results produced by the data-driven method

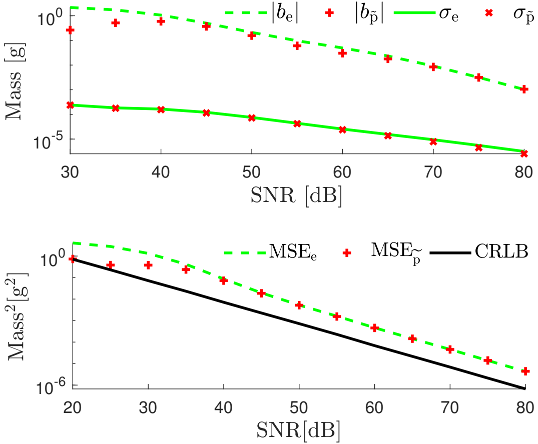


Processing a 5th order model response.



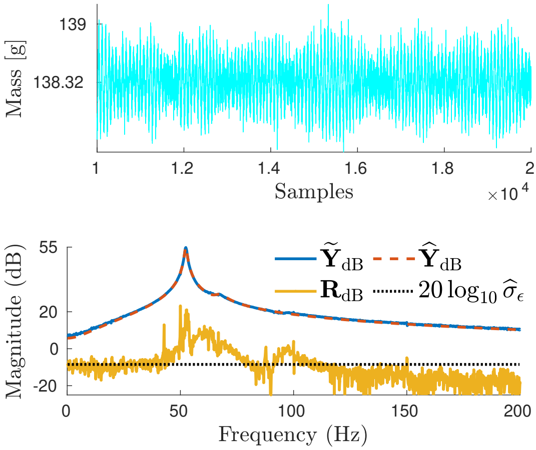
Processing a measured sensor response.

Monte Carlo simulation shows that, for SNR > 40 dB the bias and variance are well predicted



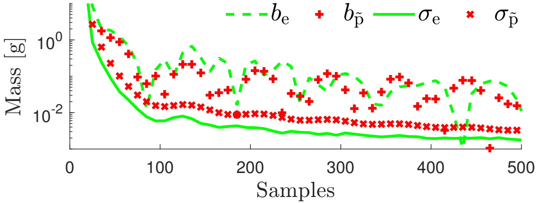
- The data-driven step input estimation MSE is at most one order of magnitude larger than the EIV problem CRLB.

There are periodic components in the measurement noise

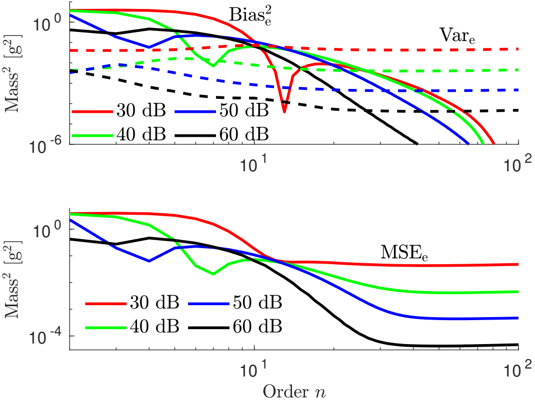


- Considering the residuals, the SNR was adjusted from 55 dB to 50 dB.

After processing 100 measured responses, the bias and variance decrease w.r.t. sample size



The estimation MSE is smallest for orders larger than the true order



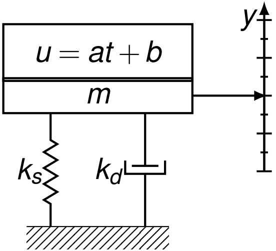
10^4 simulated responses from 5^{th} order model.

Statistical analysis and experimental validation of data-driven dynamic measurement methods

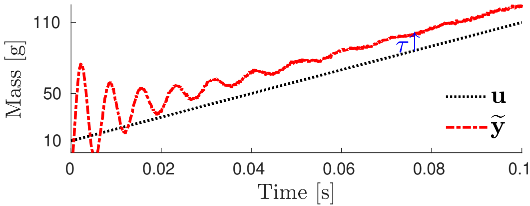
- Formulation of a data-driven step input estimation method.

- Thesis contribution.
  - Statistical analysis, structured EIV problems.
  - Experimental validation.
  - Affine input estimation methods.

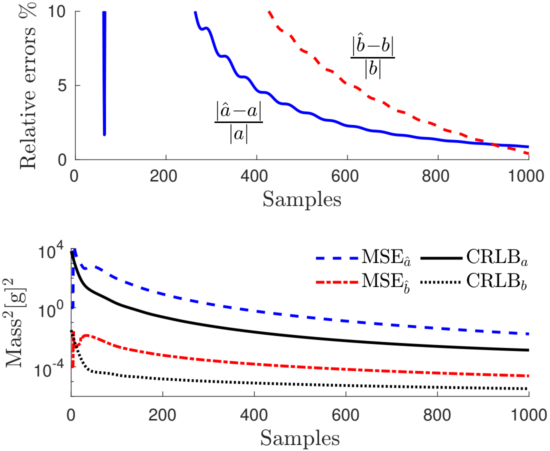
Can we estimate affine inputs with a data-driven method?



- An affine input turns this weighing system into time-varying.

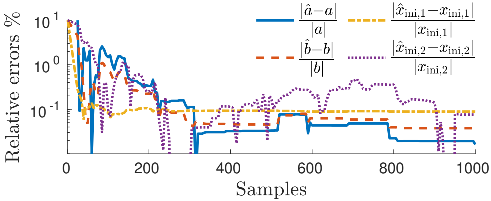


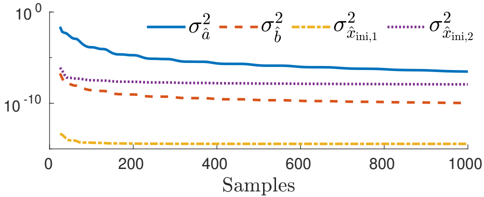
A data-driven affine input estimation method is adapted using exponentially weighted recursive LS



- Inherits statistical properties from the data-driven step input estimation method.

A maximum-likelihood affine input estimation method estimates simultaneously also the initial conditions

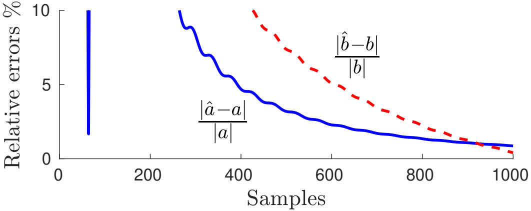




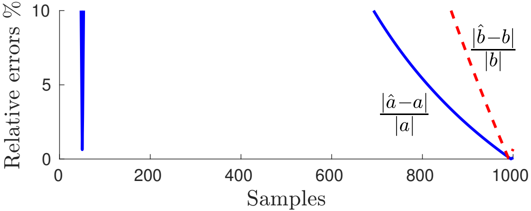
- The ML method provides the best possible estimation, but requires large computational power.

The proposed methods outperform a conventional time-varying filter

- Processing the same sensor response.
  - Data-driven method:



  - Time-varying filter:



In summary

- In the data-driven step input estimation method
  - the noise enters in the regression matrix,
  - the obtained bias and variance expressions can be used to describe the input uncertainty,
  - users can set the sample size to reach a required MSE,
  - the MSE is smallest for large model order.

Conclusions

- The data-driven input estimation methods
  - are valid for metrology applications,
  - are completed with the uncertainty assessment,
  - are robust under non Gaussian white noise,
  - can process responses from time-varying sensors.

Future work

- Design and implementation of
  - data-driven methods to estimate other input models,
  - efficient solution methods for structured EIV problems,
  - efficient online optimization methods to reduce its computational burden.

Statistical analysis and experimental validation of data-driven dynamic measurement methods

Gustavo Quintana Carapia

VUB FACULTY OF ENGINEERING