Exploiting input sparsity for joint state/input moving horizon estimation

M. Kirchner[a,b,*], J. Croes[a,b], F. Cosco[a,b], W. Desmet[a,b]

[a] *KU Leuven, Department of Mechanical Engineering, Celestijnenlaan 300, B-3001, Leuven, Belgium*
[b] *Member of Flanders Make*

Abstract

This paper proposes a novel time domain approach for joint state/input estimation of mechanical systems. The novelty consists of exploiting compressive sensing (CS) principles in a moving horizon estimator (MHE), allowing the observation of a large number of input locations given a small set of measurements. Existing techniques are characterized by intrinsic limitations when estimating multiple input locations, due to an observability decrease. Moreover, CS does not require an input to be characterized by a slow dynamics, which is a requirement of other state of the art techniques for input modeling. In the new approach, called compressive sensing–moving horizon estimator (CS-MHE), the capability of the MHE of minimizing the noise while correlating a model with measurements is enriched with an ℓ_1-norm optimization in order to promote a sparse solution for the input estimation. A numerical example shows that the CS-MHE allows for an unknown input estimation in terms of magnitude, time and location, exploiting the assumption that the input is sparse in time and space. Finally, an experimental setup is presented as validation case.

Keywords:
state estimation, input estimation, moving horizon estimation, ℓ_1-norm optimization, compressive sensing

*Corresponding author
Email address: matteo.kirchner@kuleuven.be (M. Kirchner)

1. Introduction

State estimation is a well established engineering approach aiming at recovering a complete representation of the internal condition of the system under investigation at a given time instant, and allows a system to be controlled [1]. This is achieved by correlating information coming from a model with a set of measurements. As such, it is becoming crucial in many engineering areas such as control, structural health monitoring and virtual sensing. It is well known that the Kalman filter (KF) [2] provides the optimal state estimation in case of linear systems with Gaussian, zero-mean, uncorrelated process noise [1]. However, many real world systems do not satisfy these hypothesis, and other estimation methodologies have been developed, such as the extended Kalman filter (EKF) [1] and the moving horizon estimator (MHE) [3]. The KF and its nonlinear derivations are recursive single step approaches, whereas the MHE exploits a finite length time window sliding over time. Moreover, the MHE is suited for nonlinear systems, can include constraints and has been shown to provide the correct estimation for problems with multiple optima where the EKF tends to fail [4], at the price of a higher computational cost.

All the over mentioned techniques can also combine the estimation of states and inputs. In such framework, the inputs become part of the unknowns and are referred to as augmented states [1, 5]. A joint state/input estimator is beneficial whenever the inputs are not easy to be measured or if they have a strong influence on the estimation accuracy [1, 6, 7, 8]. In fact, a joint estimator can capture the cross-coupling among all estimates by means of a single covariance matrix. However, this comes at the price of a higher computational cost and possible observability issues, which may degenerate to failure of estimation when dealing with many estimates [1]. Furthermore, additional equations are needed to model these new states. A random walk model is often employed to represent an unknown input. It is a generic approach that can be applied to different input types, but it is not suited if the number of augmented states exceeds a threshold governed by observability or if the new states are characterized by a fast dynamics [1, 6, 7, 8, 5, 9, 10, 11, 12].

Beside estimation problems, input models are key elements of force identification problems. References [13, 14, 15, 16] focus on inverse methods for force identification based on the ℓ_1-norm regularization. In particular, reference [15] describes a fast iterative shrinkage/thresholding (IST) algorithm

which leads to an accurate force reconstruction for impulses and harmonic loads. The same approach is employed in [14], where different types of shape functions are compared in order to find the best force impact representation. Furthermore, the problem of impact identification and location is solved in [13] by a two-step IST algorithm, allowing the characterization of one or two force impulses within a set of nine candidate locations. Finally, reference [16] describes the sparse deconvolution method for the reconstruction of impact forces in case of large scale ill-posed inverse problems.

The ℓ_1-norm optimization constitutes the basis of a technique called compressive sensing (CS), which is gaining attention in the fields of structural health monitoring and fault detection in order to limit the number of required sensors [17, 18, 19] and the amount of data to be transmitted for processing [20]. CS allows to acquire and recover undersampled signals, and it is based on a concept referred to as signal sparsity [21, 22, 23]. CS is currently being investigated as a powerful instrument in the framework of estimation problems based on the KF. The first example of CS as a tool to improve the KF can be found in [24], where the fact that the sparsity pattern of a signal changes slowly over time is exploited within a KF with a limited amount of measurements. This idea has been further developed in [25, 26], and two other sparsity conditions have been introduced in [27] (*i.e.*, sparsity in the state and sparsity in the innovations) in order to improve the KF performances in terms of estimation error or convergence time.

Reference [28] applies CS for the detection of a single force impact entering a mechanical system at an unknown location, such that the signal is known to be sparse in time and space, and CS allows for an accurate input estimation. Input sparsity in space has also been exploited in [29], where a frequency domain approach is proposed to identify unknown dynamic forces on a structure.

To summarize the state of the art, a joint state/input estimator is a powerful tool if the knowledge of an input is crucial to the estimation accuracy and cannot be obtained from direct measurements. However, such approach is not suitable if the number of inputs exceeds an observability constraint or a random walk is chosen to represent a high dynamics input. Recent research shows that CS can lead to an accurate force reconstruction within a static time horizon, independently from the input dynamics, provided that a set of basis functions is available to guarantee input sparsity. CS was also employed to improve a KF, proving that the sparsity within a single time step can be propagated to the next step in an iterative fashion.

This paper describes a novel joint state/input estimation approach which limits the observability issues related to the estimation of multiple variables and can be employed for the estimation of inputs characterized by a fast dynamics, thus confining the drawbacks of state augmentation with respect to observability and overcoming the limitations of the random walk model. This new methodology will be referred to as the compressive sensing–moving horizon estimator (CS-MHE). The CS-MHE is a joint state/input estimator in which an unknown input is modeled as a sparse signal. In particular, the CS-MHE exploits input sparsity in time and space. This is achieved by projecting an input onto a set of basis functions of which only a few are active. Thus, inputs distributed in time and/or space can be estimated provided that an appropriate set of basis functions is available [30].

The paper is structured as follows: first, section 2 gives an overview of the MHE and on CS. Next, section 3 illustrates the proposed CS-MHE methodology. The CS-MHE is tested first numerically in section 4 and then experimentally in section 5 for a linear time-invariant (LTI) mechanical system loaded with an impact force, proving that the CS-MHE allows to detect an unknown input on a mechanical system (*e.g.*, a force) in terms of position, magnitude and time. Finally, section 6 summarizes the conclusions.

2. An overview of MHE and CS

This section gives an overview of the two milestones on which the CS-MHE is based, *i.e.*, the moving horizon estimator (section 2.1) and compressive sensing (section 2.2).

2.1. Moving Horizon Estimator

References [4, 3, 31] describe the classical MHE for state estimation problems, which is summarized by eqs. (1–5). An estimation window of length N is defined between the discrete time steps $k = T-N+1$ and $k = T$, as shown in Fig. 1.

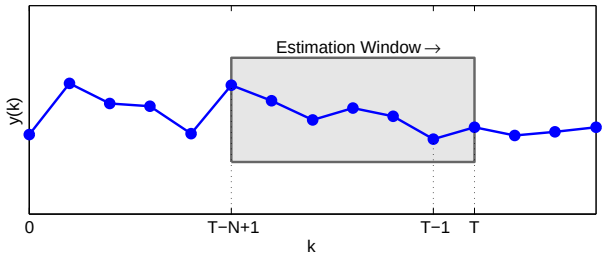

Figure 1: MHE strategy.

$$\begin{aligned}
\underset{w_a, w_k, v_k}{\text{minimize}} \quad & w_a^T P_a^{-1} w_a + \sum_{k=T-N+1}^{T-1} w_k^T Q_k^{-1} w_k + \sum_{k=T-N+1}^{T} v_k^T R_k^{-1} v_k && (1)\\
\text{subject to} \quad & x_{k+1} = f(x_k, u_k) + w_k && (2)\\
& y_k = h(x_k, u_k) + v_k && (3)\\
& x_{T-N+1} = \bar{x}_{T-N+1} + w_a && (4)\\
& x_k \in X_k, w_k \in W_k, v_k \in V_k && (5)
\end{aligned}$$

Eq. (1) is a cost function and consist of three noise terms to be minimized. From left to right, they are related to the arrival cost $w_a \in \mathbb{R}^{N_S}$, the model error $w_k \in \mathbb{R}^{N_S}$ and the measurement error $v_k \in \mathbb{R}^r$, where N_S and r are the number of states and transducers, respectively. Each variable is associated with a covariance matrix as follows:

$$\begin{aligned}
w_a &\sim \mathcal{N}(0, P_a)\\
w_k &\sim \mathcal{N}(0, Q_k)\\
v_k &\sim \mathcal{N}(0, R_k).
\end{aligned} \qquad (6)$$

Eqs. (2–3) are the state-space representation of a discretized system under the hypothesis of additive noise. Functions $f(x_k, u_k)$ and $h(x_k, u_k)$ depend on the state vector $x_k \in \mathbb{R}^{N_S}$ and on the input u_k. Eq. (4) refers to the arrival cost and finally X_k, W_k and V_k in eq. (5) are bounds on the variables x_k, w_k and v_k, respectively. Matrices Q_k and R_k have to be built according to the information that are available about the model and the measurement system. For the latter this is usually a simple task, since the accuracy of the measurement system is known, while a few assumptions are needed to choose a value for the model uncertainty.

The role of the arrival cost for the MHE is to include the past information (from $k=0$ to $k=T-N+1$) in the estimation. The arrival cost refers to the first time step of the moving window (*i.e.*, $k = T-N+1$), and it is identified by the subscript a. There are several ways to deal with the arrival cost. A first approach is simply not to take into account any information prior to the sliding window. A second method is to set a constant value for the arrival cost (*i.e.*, $P_a = constant$), and a third technique employs a recursive filter, such as the EKF or the unscented Kalman filter (UKF) [3, 32]. Furthermore, the so-called smoothed arrival cost exploits the covariance matrix of the optimization problem [33, 34, 35]. It will become clear in section 3 that the latter approach is well suited for the proposed CS-MHE, because the covariance matrix is required independently of the arrival cost. The prior information enters the estimation window through the term $\bar{x}_{T-N+1}$ in eq. 4, and its value is set according to the chosen arrival cost strategy. The effect of the arrival cost is minimal for a long window, in which a large set of data is available for the optimization, such that an extra element does not have a strong influence on the solution. For the same reason, a good estimation of the arrival cost is crucial if the window is short. The latter is the most interesting case from a practical point of view, because a short window allows for a faster computation, which may be the bottleneck for online (real-time) applications.

The optimization problem of eqs. (1–5) is solved via sequential quadratic programming (SQP), and inequality constraints can also be included. Constraints are handled depending on the chosen algorithm (*e.g.*, interior point, active set, parametric active set) [36]. When dealing with estimation problems, it is important to assign a confidence level to the estimates (states, inputs and parameters). This can be done by computing the covariance matrix of the optimization problem, which is a symmetric positive semidefinite square matrix whose diagonal entries are the variances of the optimization variables, while any nonzero off-diagonal element denotes cross-correlation between them. The covariance matrix for constrained optimization problems such as eqs. (1–5) can be calculated following the approach proposed in references [37, 38].

2.2. Compressive sensing and ℓ_1-norm optimization

Compressive sensing (CS) is a well known scheme in the field of signal and image processing [21, 39, 40]. It allows to acquire a signal at a lower sampling rate than the Nyquist-Shannon limit and subsequently to recover

a signal from an undersampled data set [23]. CS is based on a concept referred to as signal sparsity, *i.e.*, a signal can be represented (fully or in an approximate way) by just a few components (basis functions) belonging to a certain space (dictionary) [41, 22, 21]. Moreover, the sensing scheme should have a dense representation in the dictionary [22]. A simple overview on compressive sampling can be found in [21]. References [22, 42] provide the mathematical guarantees to be taken into account when dealing with compressive sampling.

Eq. (7) shows the sensing process. Vector $u \in \mathbb{R}^{\gamma}$ is an unknown signal to be measured, $y \in \mathbb{R}^{\mu}$ is a set of measurements and $\Phi \in \mathbb{R}^{\mu \times \gamma}$ is the sensing matrix, and implements the operations of signal acquisition. u is projected on the dictionary $\Psi \in \mathbb{R}^{\gamma \times \eta}$, such that $\alpha \in \mathbb{R}^{\eta}$ is a sparse representation of the input. Finally, $\Theta \in \mathbb{R}^{\mu \times \eta}$ brings together the sensing matrix and the dictionary in the so-called global sensing basis. The system is assumed to be underdetermined (typically $\mu \ll \gamma \leq \eta$) and has an infinite number of solutions. This happens for two reasons: first because the number of samples μ is kept low, and secondly because the dictionary can be overcomplete [30].

$$y = \Phi u = \Phi \Psi \alpha = \Theta \alpha \tag{7}$$

Compressive sensing can solve eq. (7), provided that y is sufficiently long and α is sufficiently sparse [22]. Specifically, the restricted isometry property (RIP) is a condition on Θ that indicates a nearly orthonormal matrix when operating on sparse vectors (such as α). If the RIP holds for a given sparsity, the rate of success of CS is very high [22, 42, 43]. Among all possible solutions of eq. (7), CS is interested in finding the sparsest one. The reason is that the number of measurements y needed to capture a sparse signal α is proportional to its sparsity (*i.e.*, to the number of nonzero elements of α). Eq. (7) can be solved through an ℓ_1-norm optimization problem, known as basis pursuit (BP) or basis pursuit denoising (BPDN, in case of solution approximation, regularization and noise filtering) [44, 45, 46, 47, 30, 22].

3. Proposed methodology: the CS-MHE

The CS-MHE formulation is given and explained here. The combined state/input estimation is achieved through a state augmentation, *i.e.*, the inputs to be estimated become part of the optimization variables. State augmentation is a standard practice in state/input/parameter estimation.

Its main drawback is that it deteriorates the observability of the system. Furthermore, extra information are required to model the newly introduced variables. This is typically done through a so-called random walk model [5], which consist on modeling an input (or paramenter) as the input (or parameter) at the previous time step plus a noise term proportional to the dynamics of that value. This approach suffers from a few limitations. First, it should be clear that such model cannot represent a sudden input change (*e.g.*, the case of an impulse). Furthermore, the number of extra variables needed may render the system unobservable. To overcome those weaknesses, CS-MHE exploits sparsity instead of a random walk to model an input.

The CS-MHE approach is given in eqs. (8–14). Most of the notation has already been explained for the classical MHE in eqs. (1–5). The new parts involve the last two terms of the cost function in eq. (8), their related bounds in eq. (14), and the new constraints denoted as eqs. (11) and (13).

$$
\begin{aligned}
\underset{w_a, w_k, v_k, \nu_{\alpha^*}, \alpha_k}{\text{minimize}} \quad & w_a^T P_a^{-1} w_a + \sum_{k=T-N+1}^{T-1} w_k^T Q_k^{-1} w_k + \sum_{k=T-N+1}^{T} v_k^T R_k^{-1} v_k \\
& + \nu_{\alpha^*}^T P_{\alpha^*}^{-1} \nu_{\alpha^*} + \lambda \sum_{k=T-N+1}^{T-1} \|\alpha_k\|_{\ell_1} && (8) \\
\text{subject to} \quad & x_{k+1} = f(x_k, u_k) + w_k && (9) \\
& y_k = h(x_k, u_k) + v_k && (10) \\
& u_k = \Psi \alpha_k && (11) \\
& x_{T-N+1} = \bar{x}_{T-N+1} + w_a && (12) \\
& \alpha^* = \bar{\alpha}^* + \nu_{\alpha^*} && (13) \\
& x_k \in X_k, w_k \in W_k, v_k \in V_k, \nu_{\alpha^*} \in N_{\alpha^*}, \alpha_k \in \Gamma_k && (14)
\end{aligned}
$$

The new CS term in eq. (8) consists of the ℓ_1-norm of the sparse representation α_k of the input. It is expressed by eq. (11), where Ψ is a sparsifying dictionary as defined in eq. (7). The optimization problem is formulated under the assumption that the input is fully unknown. If this is not the case, the equations can be easily extended to include any available input information, without loss of generality. The CS term is the only linear term of the cost function, while all other components are quadratic. A constant weight λ balances this term with the rest of the cost function, and plays a crucial role in the optimization. In fact, λ scales the contribution of the spar-

sity exploitation with regard to the noise terms of model and measurements, which are represented by the covariance matrices Q_k and R_k, respectively. Section 4.3 will give some details about the choice of λ, referring to the numerical example.

Eqs. (12–13) and their related terms in the cost function contribute to the CS-MHE by including information prior to the current window. Specifically, eq. (12) refers to the arrival cost and has already been discussed in section 2.1 for the classical MHE, while eq. (13) allows to exploit any available knowledge about an input, and will be discussed in section 3.1. Despite its similarity to a typical random walk equation, it is important to notice that eq. (13) does not refer to the input estimation. In fact, it propagates the participation factors of an already detected input to the next iteration, while the estimation is performed by the CS part. N_{α^*} and Γ_α in eq. (14) are two bounds on the newly introduced optimization variables.

3.1. Exploiting the prior information

As discussed in section 2.1, a covariance matrix can be computed to determine the accuracy of the estimates and the smoothed arrival cost. In the framework of the CS-MHE, the covariance matrix is crucial to transfer the knowledge of an input from the current window to the next time step. This is done as described in Algorithm 1 and it is implemented by eq. (13), together with its related term in eq. (8).

```
1  define window i, with k=T−N+1,...,T;
2  set optimization problem i;
3  solve optimization problem i (eqs. (8–14));
4  compute covariance i;
5  compute α*_{i|i} and n_{α* i|i};
6  compute α*_{i+1|i} and n_{α* i+1|i};
7  if n_{α* i+1|i} ≥ 1 then
8  |   compare α*_{i|i} and α*_{i|i−1};
9  |   assign P_{α* i+1|i} as follows:
10 |     P_{α* i+1|i} = P_{α* i|i} + Q_drift (for matching elements);
11 |     P_{α* i+1|i} = Q_drift (for new elements);
12 else
13 |   iteration i+1 will not have extra states;
14 end
```

Algorithm 1: Procedure for updating the sparse representation of an input.

The following list illustrates the notation of Algorithm 1 and gives all details about the propagation of the input information. Moreover, the few symbols of eqs. (8–14) that were not described yet are explained here. Items are labeled according to the line numbers they refer to in Algorithm 1.

1 Index i refers to the current optimization problem.

2–4 The optimization problem i is set and solved, and its associated covariance matrix is computed. Setting the problem includes any available prior information, *i.e.*, the arrival cost and the knowledge about a possible input. The solution of the problem returns an estimation of states and inputs. The latter are represented by the sparse vector $\alpha_{i|i}$, which is the collection of all α_k within the window[1].
Index i has been omitted from eqs. (8–14). Similarly to the notation of the arrival cost in eqs. (4) and (12), prior data are marked with a bar, such that $\bar{\alpha}^*$ corresponds to $\alpha^*_{i|i-1}$.

5 Variable $\alpha^*_{i|i}$ collects the nonzero elements of $\alpha_{i|i}$, and their number is

[1]Notation $z_{i|j}$ refers to the estimation of variable z at time step i given the information at time step j.

denoted as $n_{\alpha^* i|i}$. A nonzero element of $\alpha_{i|i}$ is considered as such if its absolute value exceeds a predefined threshold level ε_α.
The purpose of ε_α is to filter out the noise components and limit the size of the optimization problem. Its side effect is that some energy coming from the input is discarded, and consequently the input magnitude is underestimated. A simple way to avoid this energy loss will be shown in sections 4–5.
Variable $\nu_{\alpha^*} \in \mathbb{R}^{n_{\alpha^*}}$ in eqs. (8) and (13) is a noise term related to the input, and it is assumed to follow a Gaussian distribution $\nu_{\alpha^*} \sim \mathcal{N}(0, P_{\alpha^*})$, where P_{α^*} carries the weighting numbers of any estimated input.

6 Variable $\alpha^*_{i+1|i}$ of length $n_{\alpha^* i+1|i}$ transfers the current knowledge of the input to the next window. The input knowledge is shared with the next iteration $i+1$ through weighting numbers. Those can be obtained from the covariance matrix of the optimization problem. To achieve this, the set $\alpha^*_{i+1|i}$ is added to the problem as $n_{\alpha^* i+1|i}$ extra states. Note that $\alpha^*_{i|i}$ and $\alpha^*_{i+1|i}$ may differ since the window is sliding in time, and any information at $T-N+1$ is thus omitted.

7 If at least one element is shared with the next time step, the updating procedure takes place. Otherwise, nothing is transfered to the next iteration (see line 13).

8–11 The comparison between current and previous windows governs the way the input weighting numbers are updated.
In fact, if an element was present during the previous window, then the current problem and covariance matrix include its weighting number $P_{\alpha^* i|i}$. This is added to a drift term Q_{drift}, resulting in $P_{\alpha^* i+1|i}$ (line 10). On the other hand, if an element is new, only Q_{drift} is assigned to $P_{\alpha^* i+1|i}$ (line 11).

In plain text, whenever the CS-MHE detects an input, the algorithm assumes that it is highly possible that the same input will be detected also in the following iteration, until the step at which that input reaches the end of the window. In other words, the sparsity pattern of the input does not change in space, while it is shifted in time according to the sliding window. Moreover, a drift term Q_{drift} is added to the input magnitude, to relax the constraints of the next optimization problem. Note that a small Q_{drift} implies

tighter constraints, while higher values give more freedom to the solver. The choice of Q_{drift} is linked to the knowledge of the system, which is evaluated by the covariances Q_k and R_k.

The knowledge of an input is shared with the next iteration in terms of augmented states. Consequently, the number of variables of the optimization problem grows proportionally to $n_{\alpha^* i+1|i}$. The size of the problem may be kept smaller if the threshold for the input detection ε_α gets higher. The drift term to correct the input estimation can be modeled in several ways. The simplest approach employs a constant value, but a function of the time steps within the window may be more appropriate. An example will be given in sections 4–5.

4. Numerical test case

The CS-MHE described in section 3 has been tested numerically for a linear time-invariant (LTI) mechanical system. Consequently, the state-space model in eqs. (9–10) assumes the form indicated in eqs. (15–16). This section describes the numerical test case (section 4.1), discusses the observability of the system (section 4.2), suggests a method to choose the weight λ (section 4.3) and finally shows some simulation results (section 4.4).

$$x_{k+1} = Ax_k + Bu_k + w_k \tag{15}$$
$$y_k = Cx_k + Du_k + v_k \tag{16}$$

4.1. Description of the numerical test case

The cantilever beam in Fig. 2 was chosen for testing the CS-MHE methodology. Three simulated displacement transducers s_1, s_2, s_3 were located at 3 equidistant points denoted as $x(\mathrm{s}_1)$, $x(\mathrm{s}_2)$, $x(\mathrm{s}_3)$ in Table 1. The beam is modeled analytically considering a bar with uniform rectangular cross-section. A state space model was built following the procedure given in reference [28] and adapting it to a cantilever beam with displacement measurements. Geometry and material properties are summarized in Table 1. The model takes into account the first 3 eigenmodes, which are also displayed in Fig. 2. Their damping values are denoted as ζ_1, ζ_2 and ζ_3 in Table 1. The system is described by 6 states, *i.e.*, 3 position modal participation factors (MPFs) and their time derivatives [28].

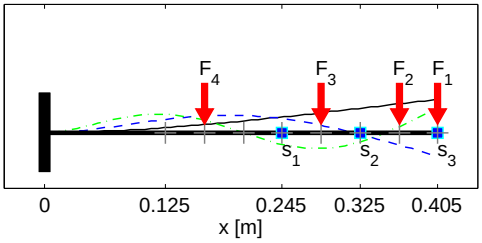

Figure 2: Numerical test case. Legend: 1st mode (——); 2nd mode (- - -); 3rd mode (- · -); spatial sampling (+); transducers (s_1, s_2, s_3); input (F_1, F_2, F_3, F_4).

Parameter	**Value**
Beam length [m]	0.405
Beam width [m]	0.025
Beam thickness h [m]	0.003
Density [kg/m^3]	7502
Young's modulus [GPa]	65.9
$x(s_1)$ [m]	0.245
$x(s_2)$ [m]	0.325
$x(s_3)$ [m]	0.405
ζ_1	0.030
ζ_2	0.037
ζ_3	0.119
ε_α [N]	1.0

Table 1: Parameters of the numerical test case.

Force ID	Time [s]	Location [m]	Magnitude [N]
F_1	0.03	0.405	-10
F_2	0.06	0.365	-5
F_3	0.09	0.285	5
F_4	0.12	0.165	10

Table 2: Location of the impacts in time and space.

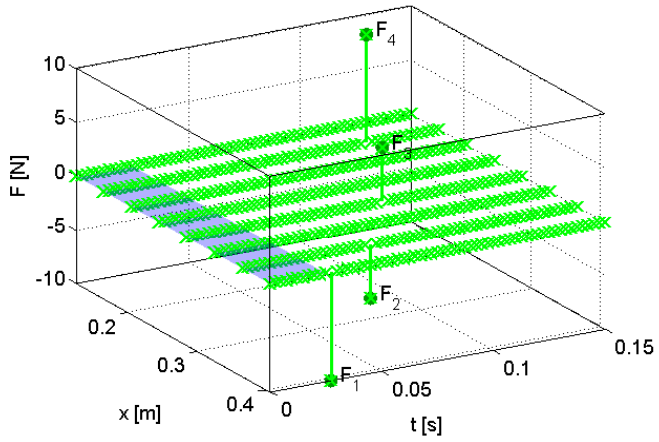

Figure 3: Reference input for the numerical test case (——×, the nonzero components are marked in a darker color), and window for the first CS-MHE iteration (light blue).

Fig. 2 shows also a spatial sampling (8 equally spaced grey crosses) and the input, which consists of 4 force impulses F_1, F_2, F_3, F_4. Those impacts enter the system at different locations and time, as indicated in Table 2 and Fig. 3. The force (green crosses) is zero except for the four impacts, such that the input signal is sparse in time and space, and there is no need to project the signal onto any specific dictionary. In other words, variables u and α are equivalent. Fig. 3 indicates the first estimation window (light blue). The system is at rest and this status does not change until F_1 is applied.

The CS-MHE can directly estimate the position of an input only if this is applied to a sampling point. If this is not the case, the input energy is spread onto the neighboring nodes. However, the exact input location can still be estimated accurately by linear interpolation, provided that the input consists of one single impulse [28]. In such context, CS outperforms the

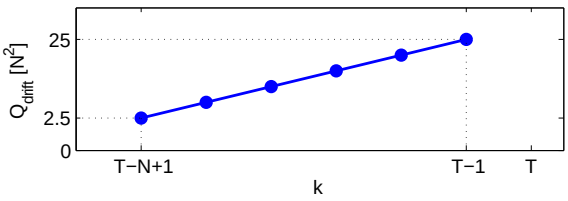

Figure 4: Q_{drift} as a linear function of the time step k.

random walk model for what the robustness in relation to the accuracy of an input location is concerned. In fact, an input applied to an unexpected location may jeopardize the estimation, since the filter does not take into account such uncertainty.

The window length has been fixed to a value of 11 time steps, such that the input estimation takes place on 10 time steps [48]. This value allows for good accuracy [3] and for a fast computation. The drift term (Q_{drift}) follows a linear function of the time step k within one window, as shown in Fig. 4. This follows the fact that the estimation is expected to be more accurate if both past and future data take part in the estimation, and this happens next to $k=T-N+1$ [49].

A model mismatch was simulated in order to investigate the influence of the modeling error. This was achieved by changing the beam thickness (h). Table 3 shows the first three eigenfrequecies of the beam for the reference test case ($h=0.003\,\text{m}$) as well as for 3 different modeling errors, characterized by a frequency mismatch indicated by $\delta_{\%}$. The choice of parameters ε_R, ε_Q and λ will become clear in section 4.3. In order to satisfy the Nyquist-Shannon sampling theorem for the highest mode, a sampling period of $2.5 \cdot 10^{-3}$ s (400 Hz) has been chosen. Note that the CS-MHE exploits compressive sampling for the observation of a large amount of input positions, and this formulation does not allow for signal acquisition with undersampled data in time.

4.2. Observability of the system

Before presenting any numerical results, let us investigate the observability of the system. Such analysis helps understanding the potentiality of the CS-MHE in comparison to joint state/input/parameter estimators that rely on a random walk model. Let us consider the LTI cantilever beam described in section 4.1 in the case in which only the states are being estimated. Then,

	$h = 0.003\,\mathrm{m}$ (reference)	$h = 0.0029\,\mathrm{m}$ ($\delta_{\%} = -3.3\,\%$)	$h = 0.0028\,\mathrm{m}$ ($\delta_{\%} = -6.7\,\%$)	$h = 0.0027\,\mathrm{m}$ ($\delta_{\%} = -10.0\,\%$)
f_1 [Hz]	8.76	8.47	8.17	7.88
f_2 [Hz]	54.88	53.05	51.22	49.39
f_3 [Hz]	153.66	148.54	143.42	138.30
ε_R [m^2]	$1.48{\cdot}10^{-8}$	$1.48{\cdot}10^{-8}$	$1.48{\cdot}10^{-8}$	$1.48{\cdot}10^{-8}$
ε_Q	$7.39{\cdot}10^{-2}$	1.48	598.74	598.74
λ	$3.27{\cdot}10^{-4}$	$2.94{\cdot}10^{-2}$	$1.98{\cdot}10^{-4}$	$3.27{\cdot}10^{-4}$

Table 3: First 3 eigenfrequencies of the beam and CS-MHE tuning parameters for the numerical test cases.

let us add one by one the possible input positions as augmented states. A zero-order random walk model is associated with each input position [5]. For the system to be observable, matrices A and C in the state-space eqs. (15–16) have to satisfy eq. (17), where $\mathcal{O}$ is the observability matrix [50, 1]. Fig. 5 shows the rank of $\mathcal{O}$ as function of the number of states N_S, which includes the 6 position and velocity MPFs as well as the augmented states for the input estimation.

$$rank(\mathcal{O}) = N_S, \qquad \text{with} \qquad \mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{N_S-1} \end{bmatrix} \tag{17}$$

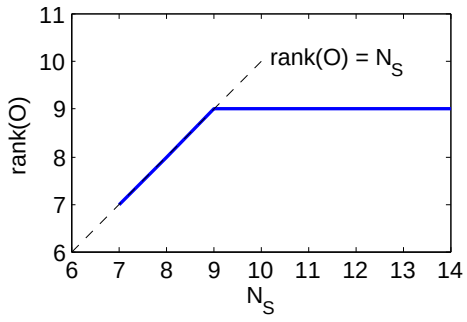

Figure 5: Rank of the observability matrix $\mathcal{O}$ as function of the number of states N_S.

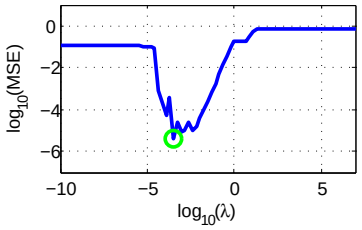

Figure 6: Choice of λ for the numerical reference test case, given a constant Q and R. MSE of the input estimation (blue line) and location of its minimum (green circle).

The graph shows that the system becomes not observable if $N_S > 9$, *i.e.*, when the number of observed positions exceeds 3. Accordingly, the numerical example in section 4.1 does not satisfy eq. (17), since the forces are applied to 4 different locations. The example is not observable if a RW is applied at each position, while it will be shown in section 4.4 that the CS-MHE is able to estimate all 4 inputs, provided that the input is sparse. The connection between sparsity and observability for a joint state/input MHE with no input model, random walk model and compressive sensing has been investigated in [48] for different scenarios. The main outcome of [48] is that the observability of the CS-MHE is linked to input sparsity, whereas observability with a random walk is constrained by the number of input locations (*i.e.*, the number of random walk models).

4.3. Choice of the balancing weight λ

The key role of the weight λ and its dependency on the model and measurement covariance matrices has already been pointed out in section 3. This crucial CS-MHE parameter is further discussed here for the numerical example. Fig. 6 shows the mean square error (MSE) of the input estimation for different values of λ, while the covariances Q and R are constant. The MSE drops within a region of λ, leading to an accurate input estimation. If λ is too small, the optimization gives more weight to the minimization of the model and measurement errors, and the sparsity of the input cannot be exploited. On the other hand, a too high value of λ would promote sparsity within a system that does not minimize any model and measurement errors, resulting in a higher MSE. An analogous simulation was carried out for different values of ε_Q and ε_R, resulting in an optimal λ for each combination, corresponding to the minimum MSE. ε_Q and ε_R are two scaling factors assigned to the covariance matrices Q and R, respectively. The study was repeated

for the other three models introduced in Table 3. Fig. 7 shows the MSE of the input estimation given as a function of ε_Q, ε_R and computed with their related optimal λ.

It is worth noticing that there are particular combinations of ε_Q and ε_R that give a smaller MSE, and those are located along a line in a log–log plane (white dashed line). This can be spotted in case of modeling error, while the reference case manifests a wider interval of low MSE. Furthermore, Fig. 8 shows the values of the optimal λ for each combination. It can be seen that they scale logarithmically while moving along the line of minimum MSE. Given a measurement setup (such that ε_R is known) and a model, there are a certain model accuracy (ε_Q) and a certain optimal λ at which the MSE has a minimum (*i.e.*, the CS-MHE performs its best). Those values are marked with a green circle and are given in Table 3.

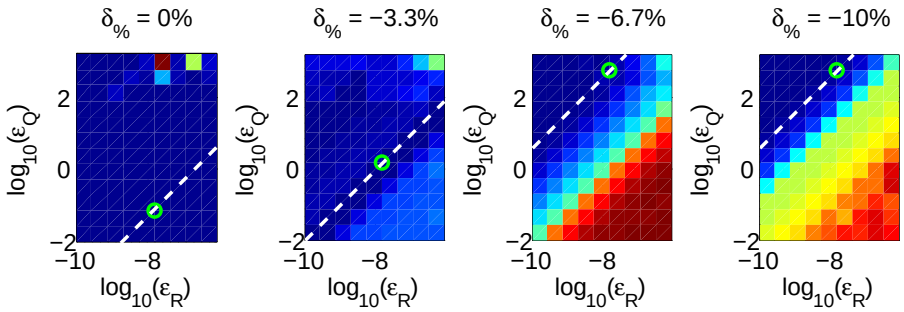

Figure 7: MSE of the input estimation as a function of ε_Q and ε_R. The white dashed lines indicate the combinations of ε_Q and ε_R that give a minimum MSE. The chosen combinations are marked with a green circle, and correspond to Table 3.

4.4. Results and discussion

This section presents the results of the numerical test case. Whenever a new measurement is available, the moving window shifts in time and performs a joint state/input estimation. Fig. 9 shows the estimation window $i = 7$ for $\delta_{\%} = 0\%$, while the whole simulation is available here. The left graphs display the states, divided into position (top) and velocity (bottom) MPFs. Their confidence intervals are also present ($\text{MPF}_n \pm 3\sigma_n$, *i.e.*, 99.7% of the normal distribution, where $n = 1, 2, 3$ identifies the first 3 eigenmodes of the structure). The right graph shows the input estimation. The time axis in Fig. 9 is relative to the current window.

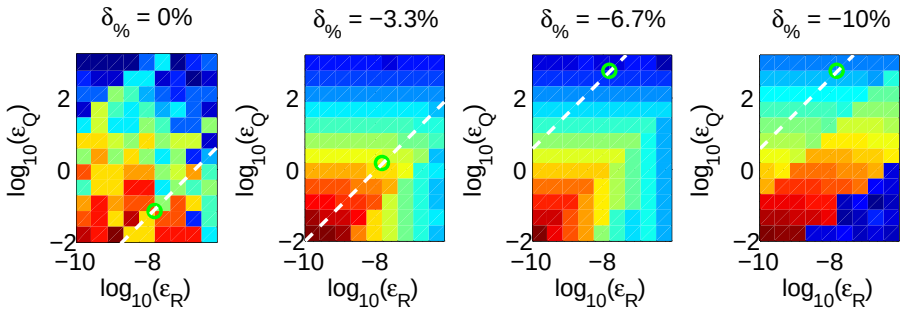

Figure 8: Optimal λ as a function of ε_Q and ε_R. The white dashed lines and the green circles corresponds to Fig. 7.

Fig. 10 shows the input estimation of the full simulation. The graph is obtained by keeping the elements $\alpha \geq \varepsilon_\alpha$ that correspond to the best estimation time step of each window, *i.e.*, $k = T - N + 1$. This follows the fact that an estimate obtained by processing both past and future data is more accurate than an estimate that relies only on prior information [49]. This approach can be adopted if the latest input estimate (*i.e.*, $k = T - 1$) is not required for specific real-time applications. Moreover, the discarded energy (due to ε_α) is evaluated and distributed to each nonzero component proportionally to its magnitude, obtaining the solid dots in Fig. 10.

Fig. 11 presents three further aspects regarding the input estimation of the full simulation, and refers to all four test cases of Table 3. The top graph shows the deviation from the reference signal for every window, expressed as MSE. It can be seen that the MSE grows proportionally with the model mismatch. Moreover, the highest values are associated with F_1. A difference in MSE is justified by the sensor positioning. A performance improvement can be obtained by optimally placing the transducers [51]. Next, the central graph shows $n_{\alpha^* i|i}$ for each window, which depends on the choice of ε_α. The dashed green line is a reference value and corresponds to the sparse signal of Fig. 3. A link between the number of inputs and the MSE can be noticed, *i.e.*, detecting more components increases the MSE. Finally, the bottom graph in Fig. 11 shows the sum of all elements in $\alpha_{i|i}$. It can be seen that the magnitudes of F_2 and F_4 do not converge to their expected values, while F_1 and F_3 do converge. The reason of this offset is in the fact that part of the model error is seen as input by the CS-MHE. This effect can be limited by a better model or a different sensor positioning which improves

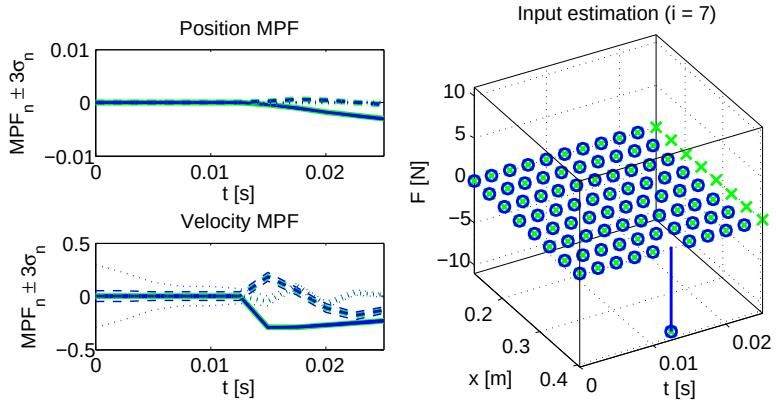

Figure 9: State/input estimation at window $i = 7$ for $\delta_{\%} = 0\%$. Legend (left graphs): 1st mode (——); 2nd mode (- - -); 3rd mode (···). The thick green lines are the reference, the thick blue lines are the CS-MHE estimation, confined into two blue thin lines that represent the confidence level ($\text{MPF}_n \pm 3\sigma_n$). Legend (right graph): reference values (——×); CS-MHE estimation (——○). The time is relative to the current window.

local observability.

As an example, let us have a look at iteration 45 of the model mismatch $\delta_{\%} = -10\%$, which is marked with a red diamond in Fig. 11. Fig. 12 shows the input estimation, where it can be seen that most of the nonzero components are located around the impulse, while some others are located away from the impulse and are due to the model mismatch. Unfortunately, it is not possible to filter out those components *a priori*. However, it can be noticed that results are accurate even in case of a high model error.

Finally, Fig. 13 (left) shows the estimation results for the case $\delta_{\%} = -10\%$ on the full time history, and all four peaks are well estimated. A way to filter out the unwanted peaks is to set a higher threshold ε_α, as shown in Fig. 13 (right). Both graphs in Fig. 13 do not include any energy correction, which can be implemented as discussed for the reference test case with $\delta_{\%} = 0\%$. An incorrect impact is estimated due to the modeling error.

To summarize, the reference test case ($\delta_{\%} = 0\%$) gave a very accurate input estimation, whereas a small error resulted when the model mismatch was increased up to $\delta_{\%} = -10\%$ on each eigenfrequency. Being the CS-MHE a model-based estimator, its performances depend on the model accuracy. The error can be reduced by improving the model and tuning the CS-MHE parameters.

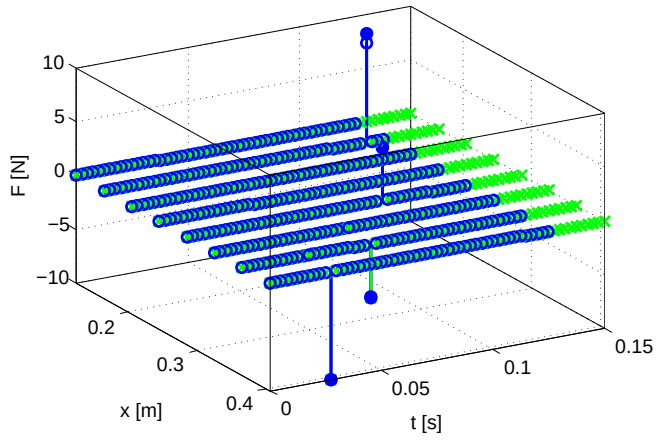

Figure 10: Global solutions at $k = T - N + 1$ and $\alpha > \varepsilon_\alpha$ for $\delta_{\%} = 0\%$. Reference (——×), CS-MHE (——○) CS-MHE corrected with the total energy (——●).

5. Experimental test case

The CS-MHE described in section 3 was applied to an experimental test case involving the cantilever beam of Fig. 14. The beam was instrumented with several transducers. First, 3 accelerometers were mounted along the beam. Moreover, a LED was placed in front of each accelerometer for vision tracking. Their location as well as the other geometry and material properties corresponds to the ones of the numerical example in section 4 (cf. Table 1). The accelerometers were used to measure the modal parameters of the beam [52], which acted as target values for updating the analytical model. The LEDs served as displacement transducers, with their motion being optically tracked by a Nikon Metrology K600 system [53].

A model updating procedure based on experimental data was performed since the presence of the accelerometers strongly influenced the dynamic behavior of the beam. Moreover, the transducers' cables generated high damping values. Table 4 shows the eigenfrequecies of the beam computed experimentally (first column) [52] and after model updating (second column). The right column indicates that the third eigenmode diverges significantly from the measurements. A better approximation could be achieved through a more detailed model such as a finite element representation [5], but an analytical model was chosen here for its simplicity, as the model type does

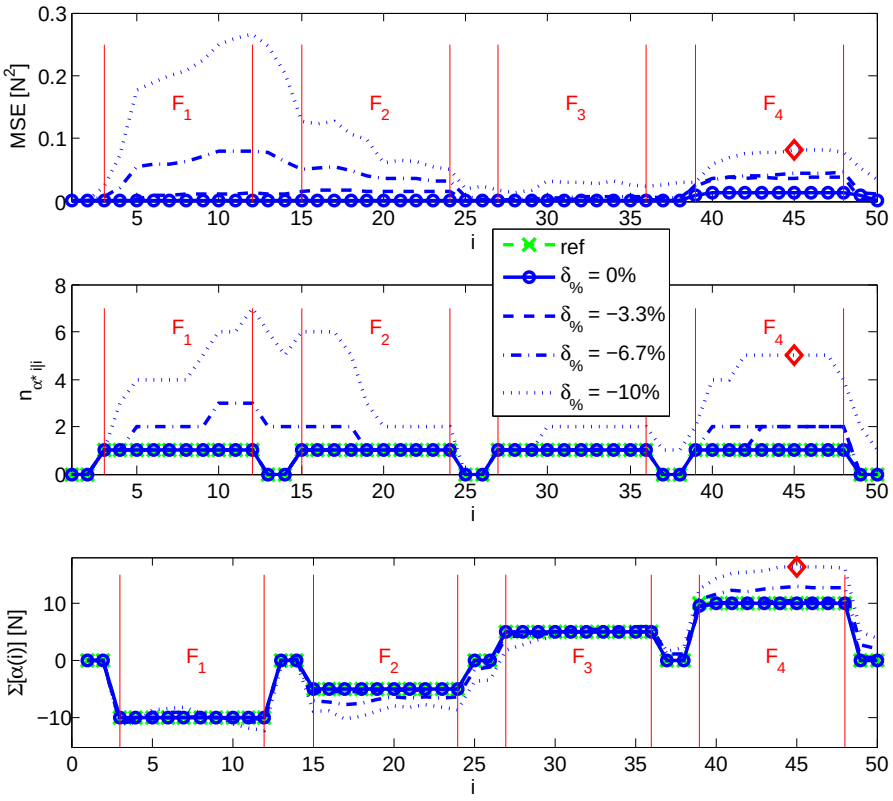

Figure 11: MSE (top), $n_{\alpha^* i|i}$ (middle) and sum of all elements in $\alpha_{i|i}$ (bottom).

not influence the proposed methodology.

A force impact F was applied by a hammer to $x(\mathrm{s}_2)$, as shown in Fig. 15. For synchronization purposes, the hammer provided a trigger signal to the data acquisition system. The exact input location in time and space was measured through a video recorded during the measurement campaign by a XIMEA MQ042CG-CM high speed camera [54] synchronized with the K600 system, which showed that the contact between the hammer and the beam holds for two time steps (the video is available here). Since no direct force measurement was available during the experiment, it has been assumed that the impulse follows a quadratic curve centered in the middle of the two nonzero components, going to zero at their adjacent time steps [55]. The four points highlighted with dark green circles in Fig. 16 belong to that curve, and serve as reference for interpreting the CS-MHE results. The magnitude

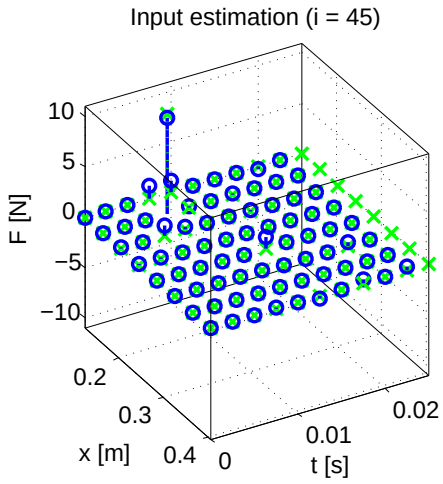

Figure 12: Input estimation at window $i = 45$ for $\delta_{\%} = -10\%$. Legend: see Fig. 9 (right).

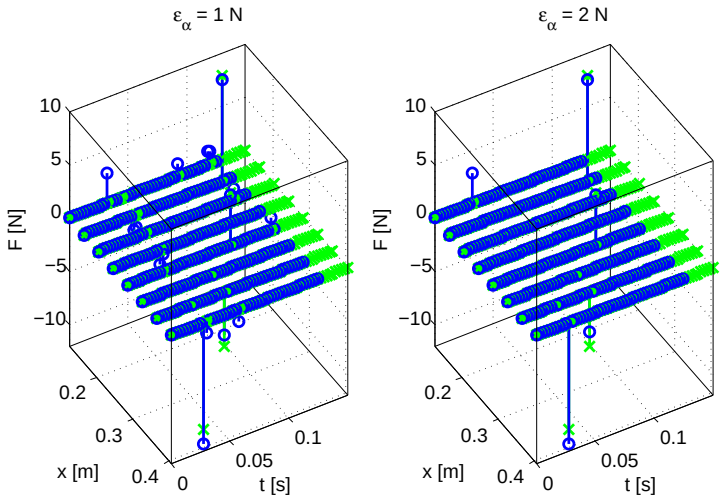

Figure 13: Input estimation for $\delta_{\%} = -10\%$ in case of $\varepsilon_\alpha = 1$N (left) and $\varepsilon_\alpha = 2$N (right). Legend: see Fig. 10.

of its peak is purely indicative and was set according to our best knowledge of the input. The values of ε_R, ε_Q and λ for the experimental test case follow from Fig. 17 and are given in the caption.

Fig. 18 shows the estimation window $i = 12$, while the whole simulation

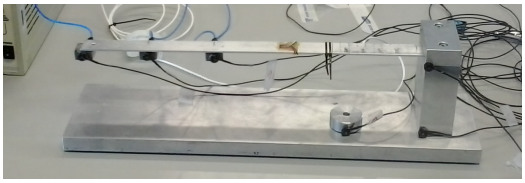

Figure 14: Experimental setup.

f_n	Experimental	Model updating	$\delta_{\%}$ [%]
f_1	8.76	8.76	0.00
f_2	61.58	54.88	-3.71
f_3	180.52	153.66	-14.88

Table 4: Comparison of the first 3 eigenfrequencies of the beam, computed experimentally and after model updating, expressed in [Hz].

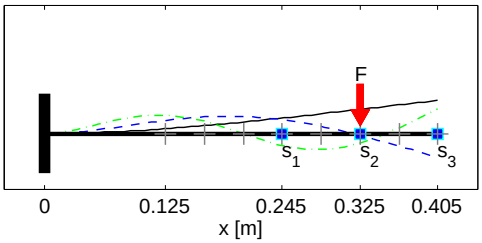

Figure 15: Experimental test case. Legend: see Fig. 2.

is available here. For the notation and the legend we refer to the numerical example in section 4.4. A few nonzero components are located in the neighborhood of the expected input position, and some of them will be filtered out since their absolute value does not exceed ε_α.

Fig. 19 shows the input estimation of the full simulation. The discarded energy (due to ε_α) is evaluated and distributed to each nonzero component proportionally to its magnitude. Three nonzero components are present, two of which are located where they are expected to be. The third one is on a neighboring node, located in time at the second nonzero component of the impact and in space at $x = 0.365$ m, *i.e.*, in the direction of the tip of the beam. A closer look at the video of the acquisition (the video is available here) shows that the hammer hits the beam in the direction of that node, which justifies the presence of the third component.

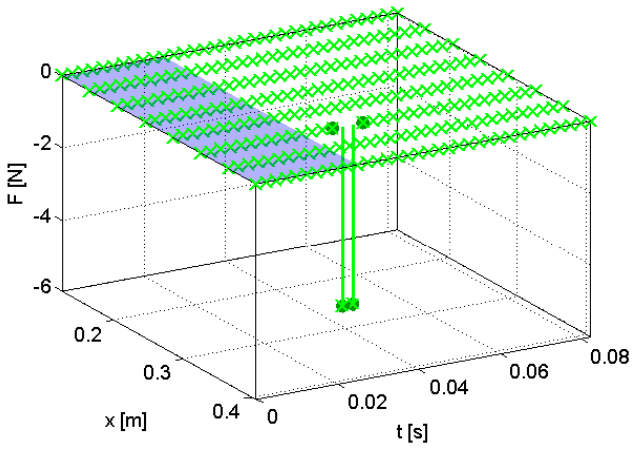

Figure 16: Reference input for the experimental test case. Legend: see Fig. 3.

Finally, Fig. 20 presents MSE and $n_{\alpha^* i|i}$ for the experimental test case. It can be seen that the estimation gets more accurate when the input approaches the end of the window, justifying the choice of Q_{drift} (cf. Fig. 4) as well as the choice of recognizing the last time step of each window as the best estimate [49]. A certain delay characterizes the input detection, revealing the intrinsic capability of a time window to detect an impulse, which is not a trivial task for estimators based on a single time step.

6. Conclusions and future work

This paper described a novel approach for combined state/input estimation based on a moving horizon estimator (MHE) enhanced with a compressive sensing (CS) term that exploits input sparsity. Such approach has been called CS-MHE and takes advantage of the MHE for correlating a model with measurements and minimizing their uncertainties in a given time window, while a ℓ_1-norm term allows estimating an input signal described by a small set of basis functions.

This latter feature is of particular interest because it allows to exploit additional known information about the input. In fact, other approaches for input estimation employ a zero order random walk model, which relies on a strong assumption concerning the dynamic range of the input, *i.e.*, the input *does not vary too much*. On the other hand, the CS-MHE requires

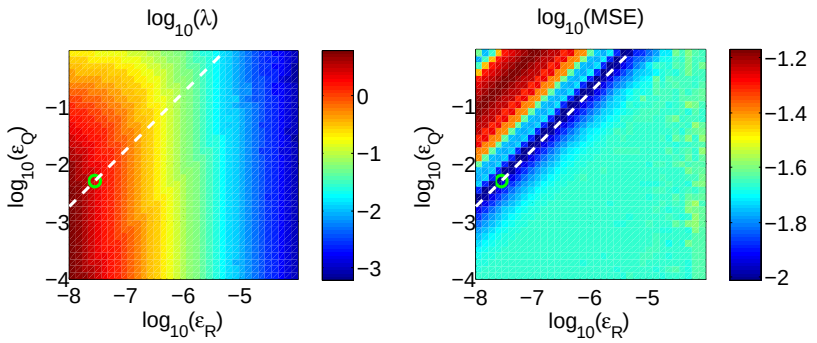

Figure 17: Optimal λ and MSE of the input estimation as functions of ε_Q and ε_R. Legend: see Figs. 7–8. The chosen values are $\varepsilon_R = 8.16 \cdot 10^{-8}$, $\varepsilon_Q = 5.10 \cdot 10^{-3}$ and $\lambda = 2.14$.

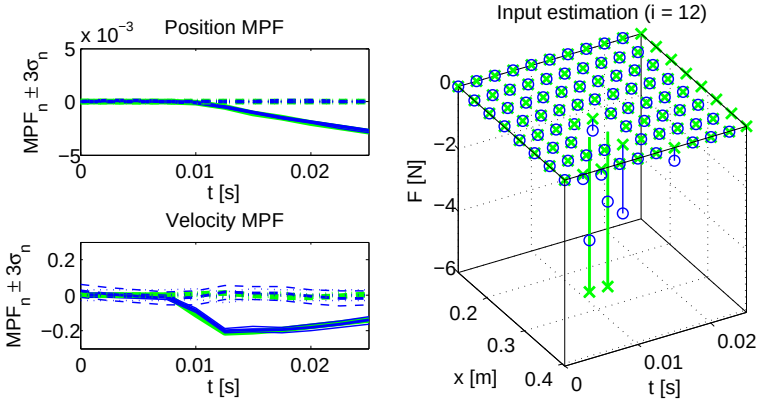

Figure 18: State/input estimation at window $i = 12$ for the experimental test case. Legend: see Fig. 9.

that the input has a sparse representation in some basis, leading to a more robust behavior in case of high dynamic ranges. Possible applications can be found in structural dynamics, structural health monitoring, process control, condition monitoring and virtual sensing.

A numerical example showed that the CS-MHE allows to estimate the states of a LTI mechanical system as well as a force impulse applied at an unknown location. The example highlighted the potential of the CE-MHE in terms of observability and input dynamic range in comparison to a random walk model. Furthermore, it exhibited the robustness of the CS-MHE with regard to modeling and measurement errors, and indicated a relationship

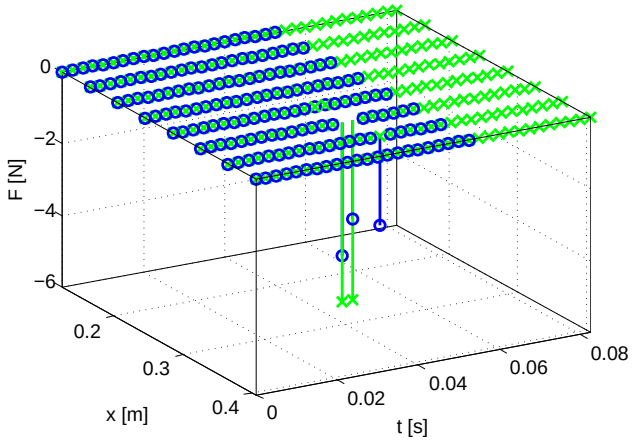

Figure 19: Global solutions at $k = T-N+1$ and $\alpha > \varepsilon_\alpha$ for the experimental test case, corrected with the total energy. Reference (——×), CS-MHE (——○).

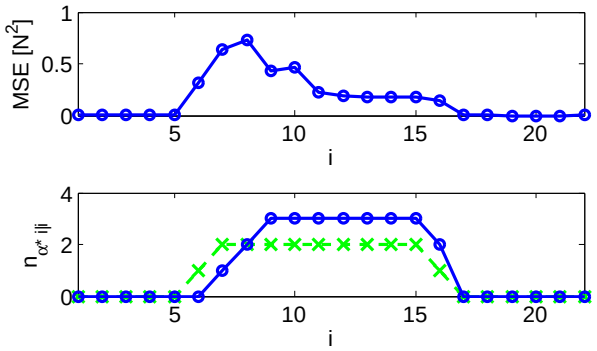

Figure 20: MSE (top) and $n_{\alpha^* i|i}$ (bottom). The dashed green line is a reference, and follows from the signal on Fig. 16.

between the CS-MHE tuning parameters. An experimental example served as validation scenario, yielding to an accurate estimation for a real system subjected to model and measurement errors.

The major drawback of the MHE over the extended Kalman filter is its higher computational effort, and this disadvantage remains also for the CS-MHE. However, new efficient algorithms are being developed for the MHE by the communities of nonlinear optimization and optimal control, from which the CS-MHE can benefit since the problem structure was not changed.

The CS-MHE relies on the choice of a few parameters, *i.e.*, the window length, the weight of the linear term of the cost function, a threshold for the input detection and a drift term to propagate the sparse representation of an input to the next iteration. This paper gave the details on how to optimally tune the weight λ given a known measurement system and a model.

This paper focused on state/input estimation, but the idea of applying compressive sensing can be applied to the wider framework of state/input/ parameter estimation. The MHE is a natural framework for nonlinear systems, thus future work will investigate more complex examples. More efficient adaptations of the proposed CS-MHE method are foreseen to handle distributed loads and multiple inputs.

The development of the CS-MHE involved an optimization problem, a few tuning parameters, an observability condition and a covariance matrix. Nevertheless, the RIP was mentioned in section 2.2 as a condition to increase the rate of success of CS. The RIP cannot be verified for arbitrary matrices, but it can be tested numerically through a Monte Carlo approach on CS problems like eq. (7), given a matrix Θ and simulating a (large) number of sparse vectors α up to a certain sparsity. Due to the more complex structure of the CS-MHE problem, such study becomes challenging. In fact, it would involve all terms of the cost function including all tuning parameters and settings of the CS-MHE, such as the number of sensors, the number of possible input locations, the number of (augmented) states, the window length and the balancing weight λ. At the same time, some of these parameters have to satisfy the observability criterion. All the numerical test cases presented along the paper as well as the experiment converge to the expected results, suggesting the success of the ℓ_1-norm optimization even if the RIP is formally not assessed. Future work will investigate this aspect.

Acknowledgments

This research was partially supported by Flanders Make, the strategic research centre for the manufacturing industry, within the MoForM project. Furthermore, the authors gratefully acknowledge the European Commission for their support of the Marie Curie program through the IAPP DEMETRA project (GA 324336). The Research Fund KU Leuven is also gratefully acknowledged for its support.

References

[1] D. Simon, Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches, Wiley-Interscience, 2006.

[2] R. E. Kalman, A new approach to linear filtering and prediction problems, Transactions of the ASME–Journal of Basic Engineering 82 (Series D) (1960) 35–45. doi:doi:10.1115/1.3662552.

[3] C. V. Rao, J. B. Rawlings, Constrained process monitoring: Moving-horizon approach, AIChE Journal 48 (1) (2002) 97–109. doi:10.1002/aic.690480111.
URL http://dx.doi.org/10.1002/aic.690480111

[4] E. L. Haseltine, J. B. Rawlings, Critical evaluation of extended kalman filtering and moving-horizon estimation, Industrial & Engineering Chemistry Research 44 (8) (2005) 2451–2460. arXiv:http://dx.doi.org/10.1021/ie0343081, doi:10.1021/ie0343081.
URL http://dx.doi.org/10.1021/ie0343081

[5] F. Naets, J. Croes, W. Desmet, An online coupled state/input/parameter estimation approach for structural dynamics, Computer Methods in Applied Mechanics and Engineering 283 (2015) 1167–1188. doi:http://dx.doi.org/10.1016/j.cma.2014.08.010.
URL http://www.sciencedirect.com/science/article/pii/S0045782514002795

[6] P. Lu, E.-J. van Kampen, C. C. de Visser, Q. Chu, Framework for state and unknown input estimation of linear time-varying systems, Automatica 73 (2016) 145 – 154. doi:http://dx.doi.org/10.1016/j.automatica.2016.07.009.
URL //www.sciencedirect.com/science/article/pii/S0005109816302771

[7] S. Gillijns, B. D. Moor, Unbiased minimum-variance input and state estimation for linear discrete-time systems, Automatica 43 (1) (2007) 111 – 116. doi:http://dx.doi.org/10.1016/j.automatica.2006.08.002.
URL http://www.sciencedirect.com/science/article/pii/S0005109806003189

[8] S. Gillijns, B. D. Moor, Unbiased minimum-variance input and state estimation for linear discrete-time systems with direct feedthrough, Automatica 43 (5) (2007) 934 – 937. doi:http://dx.doi.org/10.1016/j.automatica.2006.11.016.
URL http://www.sciencedirect.com/science/article/pii/S0005109807000222

[9] F. Naets, J. Cuadrado, W. Desmet, Stable force identification in structural dynamics using kalman filtering and dummy-measurements, Mechanical Systems and Signal Processing 50-51 (2015) 235–248. doi:http://dx.doi.org/10.1016/j.ymssp.2014.05.042.
URL http://www.sciencedirect.com/science/article/pii/S0888327014002180

[10] L. R. Ray, Nonlinear state and tire force estimation for advanced vehicle control, IEEE Transactions on Control Systems Technology 3 (1) (1995) 117–124. doi:10.1109/87.370717.

[11] L. R. RAY, Nonlinear tire force estimation and road friction identification: Simulation and experiments1,2, Automatica 33 (10) (1997) 1819 – 1833. doi:http://dx.doi.org/10.1016/S0005-1098(97)00093-9.
URL //www.sciencedirect.com/science/article/pii/S0005109897000939

[12] E. Lourens, E. Reynders, G. D. Roeck, G. Degrande, G. Lombaert, An augmented kalman filter for force identification in structural dynamics, Mechanical Systems and Signal Processing 27 (2012) 446 – 460. doi:http://dx.doi.org/10.1016/j.ymssp.2011.09.025.
URL http://www.sciencedirect.com/science/article/pii/S0888327011003931

[13] B. Qiao, X. Zhang, J. Gao, X. Chen, Impact-force sparse reconstruction from highly incomplete and inaccurate measurements, Journal of Sound and Vibration 376 (2016) 72 – 94. doi:http://dx.doi.org/10.1016/j.jsv.2016.04.040.
URL http://www.sciencedirect.com/science/article/pii/S0022460X16300955

[14] B. Qiao, Z. Mao, X. Chen, Sparse representation for the inverse problem of force identification, in: P. Sas, D. Moens, A. van de Walle (Eds.),

Proceedings of ISMA2016 including USD2016, Leuven, Belgium, 2016, pp. 1685–1696.

[15] B. Qiao, X. Zhang, C. Wang, H. Zhang, X. Chen, Sparse regularization for force identification using dictionaries, Journal of Sound and Vibration 368 (2016) 71 – 86. doi:http://dx.doi.org/10.1016/j.jsv.2016.01.030.
URL http://www.sciencedirect.com/science/article/pii/S0022460X16000596

[16] B. Qiao, X. Zhang, J. Gao, R. Liu, X. Chen, Sparse deconvolution for the large-scale ill-posed inverse problem of impact force reconstruction, Mechanical Systems and Signal Processing 83 (2017) 93 – 115. doi:http://dx.doi.org/10.1016/j.ymssp.2016.05.046.
URL http://www.sciencedirect.com/science/article/pii/S0888327016301352

[17] M. Jayawardhana, X. Zhu, R. Liyanapathirana, U. Gunawardana, Compressive sensing for efficient health monitoring and effective damage detection of structures, Mechanical Systems and Signal Processing 84, Part A (2017) 414 – 430. doi:http://dx.doi.org/10.1016/j.ymssp.2016.07.027.
URL http://www.sciencedirect.com/science/article/pii/S0888327016302503

[18] S. Bao, L. Luo, J. Mao, D. Tang, Improved fault detection and diagnosis using sparse global-local preserving projections, Journal of Process Control 47 (2016) 121 – 135. doi:http://dx.doi.org/10.1016/j.jprocont.2016.09.007.
URL http://www.sciencedirect.com/science/article/pii/S0959152416301214

[19] S. K. Perepu, A. K. Tangirala, Reconstruction of missing data using compressed sensing techniques with adaptive dictionary, Journal of Process Control 47 (2016) 175 – 190. doi:http://dx.doi.org/10.1016/j.jprocont.2016.08.008.
URL http://www.sciencedirect.com/science/article/pii/S0959152416301093

[20] D. Mascareas, A. Cattaneo, J. Theiler, C. Farrar, Compressed sensing techniques for detecting damage in structures, Structural Health Monitoring 12 (4) (2013) 325–338. arXiv:http://dx.doi.org/10.1177/1475921713486164, doi:10.1177/1475921713486164.
URL http://dx.doi.org/10.1177/1475921713486164

[21] B. Hayes, The best bits, American Scientist 97 (4) (2009) 276–280. doi:10.1511/2009.79.276.

[22] E. Candes, M. Wakin, An introduction to compressive sampling, Signal Processing Magazine, IEEE 25 (2) (2008) 21–30. doi:10.1109/MSP.2007.914731.

[23] S. Qaisar, R. Bilal, W. Iqbal, M. Naureen, S. Lee, Compressive sensing: From theory to applications, a survey, Journal of Communications and Networks 15 (5) (2013) 443–456. doi:10.1109/JCN.2013.000083.

[24] N. Vaswani, Kalman filtered compressed sensing, in: Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on, 2008, pp. 893–896. doi:10.1109/ICIP.2008.4711899.

[25] A. Carmi, P. Gurfil, D. Kanevsky, Methods for sparse signal recovery using kalman filtering with embedded pseudo-measurement norms and quasi-norms, Signal Processing, IEEE Transactions on 58 (4) (2010) 2405–2409. doi:10.1109/TSP.2009.2038959.

[26] D. Kanevsky, A. Carmi, L. Horesh, P. Gurfil, B. Ramabhadran, T. Sainath, Kalman filtering for compressed sensing, in: Information Fusion (FUSION), 2010 13th Conference on, 2010, pp. 1–8. doi:10.1109/ICIF.2010.5711877.

[27] A. Charles, M. Asif, J. Romberg, C. Rozell, Sparsity penalties in dynamical system estimation, in: Information Sciences and Systems (CISS), 2011 45th Annual Conference on, 2011, pp. 1–6. doi:10.1109/CISS.2011.5766179.

[28] D. Ginsberg, C.-P. Fritzen, New approach for impact detection by finding sparse solution, in: P. Sas, D. Moens, H. Denayer (Eds.), Proceedings of ISMA2014 including USD2014, Leuven, Belgium, 2014, pp. (2014), pp. 2043–2056.

[29] A. Rezayat, V. Nassiri, B. D. Pauw, J. Ertveldt, S. Vanlanduit, P. Guillaume, Identification of dynamic forces using group-sparsity in frequency domain, Mechanical Systems and Signal Processing 70-71 (2016) 756 – 768. doi:http://dx.doi.org/10.1016/j.ymssp.2015.09.015.
URL http://www.sciencedirect.com/science/article/pii/S0888327015004100

[30] S. S. Chen, D. L. Donoho, M. A. Saunders, Atomic decomposition by basis pursuit, SIAM Rev. 43 (1) (2001) 129–159. doi:10.1137/S003614450037906X.
URL http://dx.doi.org/10.1137/S003614450037906X

[31] C. V. Rao, Moving horizon strategies for the constrained monitoring and control of nonlinear discrete-time systems, Ph.D. thesis, University of Wisconsin-Madison (February 2000).
URL http://jbrwww.che.wisc.edu/

[32] C. C. Qu, J. Hahn, Computation of arrival cost for moving horizon estimation via unscented kalman filtering, Journal of Process Control 19 (2) (2009) 358 – 363. doi:http://dx.doi.org/10.1016/j.jprocont.2008.04.005.
URL http://www.sciencedirect.com/science/article/pii/S0959152408000711

[33] C. V. Rao, J. B. Rawlings, J. H. Lee, Constrained linear state estimation–a moving horizon approach, Automatica 37 (10) (2001) 1619 – 1628. doi:http://dx.doi.org/10.1016/S0005-1098(01)00115-7.
URL http://www.sciencedirect.com/science/article/pii/S0005109801001157

[34] R. Lopez Negrete de la Fuente, Nonlinear programming sensitivity based methods for constrained state estimation, Ph.D. thesis, Carnegie Mellon University (Pittsburgh, PA), Department of Chemical Engineering (2011).
URL http://repository.cmu.edu/dissertations/174

[35] R. Lopez Negrete de la Fuente, B. T. Lorenz, A moving horizon estimator for processes with multi-rate measurements: A nonlinear programming sensitivity approach, Journal of Process Control 22 (4) (2012) 677 – 688. doi:http://dx.doi.org/10.1016/j.jprocont.2012.01.013.

URL http://www.sciencedirect.com/science/article/pii/S0959152412000169

[36] H. Ferreau, Model predictive control algorithms for applications with millisecond timescales, Ph.D. thesis, K.U. Leuven (2011).

[37] H. G. Bock, E. Kostina, O. Kostyukova, Covariance matrices for parameter estimates of constrained parameter estimation problems, SIAM Journal on Matrix Analysis and Applications 29 (2) (2007) 626–642. arXiv:http://dx.doi.org/10.1137/040617893, doi:10.1137/040617893.
URL http://dx.doi.org/10.1137/040617893

[38] H. Bock, S. Körkel, E. Kostina, J. Schlöder, Robustness aspects in parameter estimation, optimal design of experiments and optimal control, in: Reactive Flows, Diffusion and Transport, Springer Berlin Heidelberg, 2007, pp. pp. 117–146. doi:10.1007/978-3-540-28396-6_6.
URL http://dx.doi.org/10.1007/978-3-540-28396-6_6

[39] M. Elad, M. Figueiredo, Y. Ma, On the role of sparse and redundant representations in image processing, Proceedings of the IEEE 98 (6) (2010) 972–982. doi:10.1109/JPROC.2009.2037655.

[40] M. Elad, Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing, 1st Edition, Springer Publishing Company, Incorporated, 2010.

[41] R. Baraniuk, Compressive sensing [lecture notes], Signal Processing Magazine, IEEE 24 (4) (2007) 118–121. doi:10.1109/MSP.2007.4286571.

[42] E. J. Candes, The restricted isometry property and its implications for compressed sensing, Comptes Rendus Mathematique 346 (9-10) (2008) 589–592. doi:http://dx.doi.org/10.1016/j.crma.2008.03.014.
URL http://www.sciencedirect.com/science/article/pii/S1631073X08000964

[43] M. Kirchner, E. Nijman, Cylindrical nearfield acoustical holography using compressive sampling: feasibility and numerical examples, in: P. Sas, D. Moens, H. Denayer (Eds.), Proceedings of ISMA2014 including USD2014, Leuven, Belgium, 2014, pp. 1531–1546.

[44] E. Candes, J. Romberg, T. Tao, Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information, Information Theory, IEEE Transactions on 52 (2) (2006) 489–509. doi:10.1109/TIT.2005.862083.

[45] E. Candes, T. Tao, Decoding by linear programming, Information Theory, IEEE Transactions on 51 (12) (2005) 4203–4215. doi:10.1109/TIT.2005.858979.

[46] D. Donoho, Compressed sensing, Information Theory, IEEE Transactions on 52 (4) (2006) 1289–1306. doi:10.1109/TIT.2006.871582.

[47] D. Donoho, J. Tanner, Thresholds for the recovery of sparse solutions via l1 minimization, in: Information Sciences and Systems, 2006 40th Annual Conference on, 2006, pp. 202–206. doi:10.1109/CISS.2006.286462.

[48] M. Kirchner, J. Croes, F. Cosco, B. Pluymers, W. Desmet, Compressive sensing-moving horizon estimator for combined state/input estimation: an observability study, in: P. Sas, D. Moens, A. van de Walle (Eds.), Proceedings of ISMA2016 including USD2016, Leuven, Belgium, 2016, pp. 2947–2962.

[49] M. Sain, J. Massey, Invertibility of linear time-invariant dynamical systems, Automatic Control, IEEE Transactions on 14 (2) (1969) 141–149. doi:10.1109/TAC.1969.1099133.

[50] S. Korovin, V. Fomichev, State Observers for Linear Systems with Uncertainty, De Gruyter Expositions in Mathematics, No. 51, De Gruyter, Berlin, Boston, 2009.
URL http://www.degruyter.com/view/product/176715

[51] E. M. Hernandez, Efficient sensor placement for state estimation in structural dynamics, Mechanical Systems and Signal Processing 85 (2017) 789 – 800. doi:http://dx.doi.org/10.1016/j.ymssp.2016.09.005.
URL //www.sciencedirect.com/science/article/pii/S0888327016303442

[52] B. Peeters, H. Van der Auweraer, P. Guillaume, J. Leuridan, The polymax frequency-domain method: A new standard for modal parameter estimation?, Shock and Vibration 11 (3-4) (2004) 395–409. doi:10.1155/2004/523692.

[53] Nikon Metrology NV, http://www.nikonmetrology.com.
URL http://www.nikonmetrology.com

[54] XIMEA, https://www.ximea.com.
URL https://www.ximea.com

[55] K. L. Johnson, Contact Mechanics, Cambridge University Press, 1985, cambridge Books Online.
URL http://dx.doi.org/10.1017/CBO9781139171731