Joint state/input estimation with a Fourier dictionary for the input representation: effect of spectral leakage

M Kirchner$^{1,2}$, J Croes$^{1,2}$ and W Desmet$^{1,2}$

$^1$ KU Leuven, Department of Mechanical Engineering, Celestijnenlaan 300, B-3001 Leuven, Belgium.
$^2$ DMMS lab, Flanders Make.

E-mail: matteo.kirchner@kuleuven.be

Abstract. The compressive sensing-moving horizon estimator (CS-MHE) is an approach for joint state/input estimation. It integrates compressive sensing principles into a moving horizon estimator, enabling to exploit shape functions to model an input, resulting in better observability and wider input bandwidth in comparison to other input models. With the final aim of using the CS-MHE for the estimation of forces and torques in rotating machinery which exhibit some form of periodicity, the authors have recently investigated Fourier shape functions. A first experimental validation showed very accurate estimation under the hypothesis of no spectral leakage, i.e., the MHE window, the sampling rate and the Fourier dictionary match a known input periodicity, with the input consisting of few sinusoidal components. This paper discusses the problem of spectral leakage that can result if the MHE window does not match the signal periodicity. In particular, we show how to remove the link between the MHE window and a Fourier dictionary, we discuss how the autocorrelation can be employed to detect the periodicity and we list further possible alternatives to enhance a sparse solution. The discussion is supported by numerical and experimental investigations.

1. Introduction

The compressive sensing-moving horizon estimator (CS-MHE) is a multistep time domain approach for joint state/input estimation [1, 2, 3, 4]. It exploits a moving horizon estimator (MHE) that minimizes the noise while correlating a model with measurements [5, 6], together with compressive sensing (CS) principles [7], which enable the observation of a relatively large number of input locations for a small set of measurements. Other state of the art estimation techniques are characterized by intrinsic limitations when estimating multiple input locations, due to an observability decrease [2]. The choice of compressive sensing to represent an input [8, 9, 10, 11, 12, 13] constitutes an alternative to the widely employed random walk model [14, 15]. In comparison to the latter, CS allows to efficiently model a high input bandwidth with respect to the sampling rate by exploiting a known input shape, resulting in the need for a coarser sampling [4].

We initially developed the CS-MHE to jointly estimate the states and a force impulse on a mechanical system in terms of input magnitude, time and position, under the assumption of a sparse input [1]. Furthermore, we formulated the CS-MHE to estimate complex quantities such
as Fourier components, which can be instrumental to model a periodic load [3, 4]. As an example, we mention forces and torques in rotating machines, whose characteristics are of quasi-periodic nature [3]. Estimating forces and torques is of paramount importance for condition monitoring and control engineering. In fact, these physical quantities are difficult or even impossible to be measured in a cheap, accurate and non intrusive manner, while their knowledge can improve a product or a process [4]. When dealing with Fourier basis functions, unwanted frequency components may appear due to spectral leakage [16]. This can be the result of the sampling scheme and/or it can be due to the finite portion of signal taken into account. In the context of the CS-MHE, a proper choice of the sampling rate and of the length of an estimation window are crucial to enhance sparsity, thus improving observability and estimation accuracy [2].

In this paper we report the formulation of the CS-MHE, focusing on the structure of Fourier basis functions (section 2). This allows to reveal the link between the MHE window and the Fourier dictionary, and to propose a few possible approaches to modify this link in order to limit the effect of spectral leakage. Furthermore, we focus on the use of the autocorrelation in order to detect the input periodicity (section 3). We support the analysis by some experiments that involve a structure excited by a shaker with a periodic force profile. Finally, we summarize the results (section 4).

2. The CS-MHE with a Fourier dictionary to represent a periodic input

This section gives the formulation of the CS-MHE for the case of a periodic input described by Fourier components. The methodology was already proposed in [3] and tested experimentally in [4]. Here we give further details on the structure of the Fourier basis functions in section 2.1, and we outline the CS-MHE formulas in section 2.2.

2.1. Fourier dictionary

The basic concept of compressive sensing is signal sparsity, i.e., the signal to be reconstructed (a force input in our case) must have only few nonzero elements [7]. A generic periodic signal in time domain does not comply with this requirement, whereas its Fourier transform may consist of only few active components. Consequently, we make use of a Fourier dictionary to represent a periodic input as a sparse signal.

Under the assumptions of regular sampling and equal amount of samples and Fourier components, a Fourier dictionary $\Psi$ is a square matrix that applies the inverse Fourier transform to a signal, i.e., $u = \Psi \alpha$, where vectors $u$ and $\alpha$ are a force input and its sparse Fourier representation, respectively [3, 17]. Fourier components form an orthogonal set (complete dictionary), i.e., each component cannot be represented by a linear combination of other components [18]. Eq. (1) shows that $\Psi$ is made of elements $\psi_{k,j}$ at row $k$ and column $j$. Subscripts $k$ and $j$ refer to a sampling point and to a Fourier component, respectively. Each column of $\Psi$ is thus a Fourier basis function of the dictionary. For the CS-MHE formulation presented in [1, 3, 4], $\Psi$ has size $(N-1) \times (N-1)$, where $N$ is the number of time steps that define the signal length, as illustrated in Fig. 1.

$$
\Psi = \begin{bmatrix}
\psi_{T-N+1,T-N+1} & \cdots & \psi_{T-N+1,T-1} \\
\vdots & \ddots & \vdots \\
\psi_{T-1,T-N+1} & \cdots & \psi_{T-1,T-1}
\end{bmatrix}
$$

The elements of $\Psi$ are built from the formulas of the inverse discrete Fourier transform (DFT), as indicated in eq. (2) for the case of a one-dimensional signal, where $i$ indicates the imaginary part of a complex number. The base length $L_0 = dt (N-1)$ of the Fourier shape functions is linked to the length of the signal ($N$) and to the sampling scheme, being $dt$ the sampling
Such dictionary avoids spectral leakage only if all Fourier components of the input fit exactly in the base length $L_0$. Unfortunately, for some applications it is not possible to set the acquisition scheme and signal length such that they comply with these requirements, and spectral leakage occurs. This jeopardizes sparsity and possibly introduces observability issues [2]. In section 3 we will discuss how to cope with this issues. Finally, it is worth mentioning that a dictionary for compressive sensing based on Fourier basis functions can also refer to a non uniform sampling, and this may result in a rectangular overcomplete dictionary [18], where $k$ and $j$ in are disconnected from each other.

$$\psi_{k,j} = \frac{1}{\sqrt{N-1}} e^{\frac{2\pi i}{N-1} j(k-\frac{1}{2})} \tag{2}$$

$$= \sqrt{\frac{dt}{L_0}} e^{\frac{2\pi i}{L_0} j(k-\frac{1}{2})} \tag{3}$$

2.2. Formulation of the CS-MHE with a Fourier dictionary

This section gives the formulation of the CS-MHE for the case of an input described by Fourier components. Further details can be found in [3, 4]. The formulas consider a linear discrete time state-space system such as eqs. (4–5), where $x_k$ is the state vector, and the input $u_k$ is represented by a linear superposition of Fourier basis functions $\psi_j$ multiplied by a scale factor $\alpha_k$. Eqs. (4–5) can also be obtained by linearizing a nonlinear system. Subscript $k$ denotes the current time step. 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. $w_k$ and $v_k$ are two noise terms associated to model and measurements, respectively.

$$x_{k+1} = A_k x_k + B_k u_k + w_k = A_k x_k + B_k \sum_{j=T-N+1}^{T-1} \psi_{k,j} \alpha_k + w_k \tag{4}$$

$$y_k = C_k x_k + D_k u_k + v_k = C_k x_k + D_k \sum_{j=T-N+1}^{T-1} \psi_{k,j} \alpha_k + v_k \tag{5}$$

Under the hypothesis of additive noise (which can always be satisfied thanks to discretization [15, 19]), the noise terms can be written down explicitly and substituted in the CS-MHE optimization problem, that we present in eqs. (6–7) [2, 3, 4]. Eq. (6) is the cost function of the optimization problem that the CS-MHE solves at every time step, and contains five terms to be minimized, which are weighted through covariance matrices. From top to bottom, they refer to the arrival cost $P_a \in \mathbb{R}^{N_S \times N_S}$, the model error $Q_k \in \mathbb{R}^{N_S \times N_S}$ and the measurement error $R_k \in \mathbb{R}^{n_r \times n_r}$, where $N_S$ and $n_r$ are the number of states and transducers, respectively.
\(Q_k\) and \(R_k\) are related to the noise terms in eqs. (4–5), being \(w_k \sim \mathcal{N}(0, Q_k)\) and \(v_k \sim \mathcal{N}(0, R_k)\). Furthermore, covariance \(P_a\) weights the propagation of any estimated input to the next time step [1], and finally the \(\ell_1\)-norm minimization of the sparse representation \(\alpha_k\) of the input is balanced with the other components of the cost function through a constant weight \(\lambda\), i.e., \(\lambda\) balances the noise on model and measurements with the input sparsity. In fact, the \(\ell_1\)-norm minimization promotes a sparse solution [20]. For the details of compressive sensing and \(\ell_1\)-norm minimization we refer to [7] and references therein.

\[P_a\] and \(P_\alpha\) in eq. (6) weight any information prior to the current estimation window, which is marked with a bar (\(\bar{x}_{T-N+1}\) for the arrival cost and \(\bar{\alpha}\) for any inputs). The arrival cost carries the information from \(k = 0\) to \(k = T-N+1\). We computed it following a smoothing approach, exploiting the covariance matrix of the optimization problem [21, 22, 23]. Finally, eq. (7) contains some bounds on the optimization variables (superscripts \(\text{LB}\) and \(\text{UB}\) denote lower and upper bounds, respectively) [4].

\[
\min_{x_k, \alpha_k} \left( x_{T-N+1} - \bar{x}_{T-N+1} \right)^\top P_a^{-1} \left( x_{T-N+1} - \bar{x}_{T-N+1} \right) \\
+ \sum_{k=T-N+1}^{T-1} \left( x_{k+1} - A_k x_k - B_k \sum_{j=T-N+1}^{T-1} \psi_{k,j} \alpha_k \right)^\top Q_k^{-1} \left( x_{k+1} - A_k x_k - B_k \sum_{j=T-N+1}^{T-1} \psi_{k,j} \alpha_k \right) \\
+ \sum_{k=T-N+1}^{T} \left( y_k - C_k x_k - D_k \sum_{j=T-N+1}^{T-1} \psi_{k,j} \alpha_k \right)^\top R_k^{-1} \left( y_k - C_k x_k - D_k \sum_{j=T-N+1}^{T-1} \psi_{k,j} \alpha_k \right) \\
+ (\alpha - \bar{\alpha})^\top P_a^{-1} (\alpha - \bar{\alpha}) + \lambda \sum_{k=T-N+1}^{T-1} \|\alpha_k\|_{\ell_1} 
\]

subject to \(x \in [x_{\text{LB}}, x_{\text{UB}}], \alpha \in [\alpha_{\text{LB}}, \alpha_{\text{UB}}]\) (6)

The optimization problem (6–7) can be written in matrix form. It includes complex variables and it can be recast into a standard form second order cone problem (SOCP) [3, 4, 24, 25]. We solved the SOCP in MATLAB\textsuperscript{®} by exploiting the modeling language YALMIP [26] and the solver MOSEK [27].

3. Spectral leakage

Up to here we mainly reported the state of the art of the CS-MHE with a Fourier dictionary, and we added a few details concerning the structure of the dictionary \(\Psi\). During previous numerical as well as experimental examples we showed that the CS-MHE generates very accurate estimates under the test settings which we chose, i.e., we worked with sinusoidal components for which an exact number of cycles is present in one estimation window and the sampling rate is an exact multiple of the components [3, 4]. This promotes sparsity since it avoids any spectral leakage. In this section we discuss different ways to detect the periodicity of a signal, aiming at minimizing spectral leakage in the generic case in which periodicity is not known \textit{a priori}. 

\[
\]
Starting from eqs. (2–3), we consider a few ways to adapt the dictionary in order to match the detected periodicity, going in the direction of dictionary learning [28, 29, 30]. If the input signal under investigation would be available (instead of being the object of an estimation problem), we could take a long portion of it and perform a DFT (standard, fast, windowed [16], or following a specific algorithm such as the ones proposed in [31, 32]). Unfortunately we are dealing with an estimation problem, and we cannot rely on the availability of a signal. Moreover, a DFT applied to a short signal would not generate reliable results (an estimation window is usually short in order to limit the computational cost). In case of rotating machinery, the easiest approach to detect the periodicity is a direct measurement of the rotation angle. In fact, we expect periodicity to be a multiple of a full rotation. An optical encoder or a proximity sensor can detect the rotation period, and this information can be used to tune the base length of the Fourier dictionary. In case a direct measurement is not possible, we propose to operate in time domain. Specifically, we focus our analysis on the autocorrelation [33], which we will introduce and discuss in section 3.2.

Independently of the methodology to detect the base period of the load, we can modify the base length of the dictionary following different approaches, which can sometimes be considered simultaneously. In the list that follows we outline four methodologies, which correspond to the graphs in Fig. 2.

(i) The estimation window as well as the number of basis functions do not change ($N$ stays constant), whereas the base length of the dictionary is adapted according to the detected input periodicity: this is the simplest approach to implement, since the structure of the dictionary does not change. In fact, only $L_0$ in eq. 3 is updated in response to a detected periodicity (first graph in Fig. 2). In this paper we focus on this approach, where the sizes of the dictionary and of the MHE window remain constant.

(ii) The length of the estimation window $N$ changes together with the number of basis functions, i.e., the dictionary remains a square matrix according to the new window length: modifying the MHE window length implies a different problem size (governed by $N$) and consequently a different computational effort (second graph in Fig. 2). It may be worth to apply such philosophy if the window needs to be shortened, but on the other hand this is not appealing in case the window has to be increased, since the computational cost would also grow. In parallel to this, the considerations in the previous point still apply, and we can see this approach as a possible way to lower the computational cost at the price of some effort in adapting the problem size.

(iii) The dictionary and the MHE window length are completely disconnected, $\Psi$ may become a rectangular matrix, and possibly the sampling scheme changes in order to reduce the computational cost: a Fourier dictionary can also refer to an irregular and/or undersampled acquisition scheme (in comparison with the Nyquist-Shannon sampling theorem). Such case is particularly interesting for compressive sensing [3], and cannot be obtained by simply inverting the DFT matrix (third graph in Fig. 2). References [17, 34] provide further details and examples. In the framework of the CS-MHE we did not examine this procedure yet, since it requires quite some modifications in the implementation of the dictionary in the estimator. Furthermore, we are currently focusing on the estimation accuracy rather than on computational performance. The basic principle is to (randomly) omit some samples in order to reduce the computational load. This would allow to take into account a relatively long time window by processing less sampling points, decreasing the computational load.

(iv) Build an ad-hoc dictionary: it is possible that a force acting on a mechanical component is influenced by strong dynamical effects which are independent of the periodic nature of an external load. Consequently, a Fourier dictionary based on a single base frequency may not constitute a satisfactory approximation. It is then possible to build an ad-hoc
dictionary that includes the eigenfrequencies of the system as well as the harmonic load. Such dictionary may be overcomplete [18], may or may not have a random sampling scheme, and it certainly is an interesting open point for future research (fourth graph in Fig. 2).

The common element in each of the four proposed approaches is the fact that the MHE window length can be disconnected from the base length that characterizes a Fourier series. Besides the purpose of modifying the dictionary to limit spectral leakage, this fact can assume importance in case we want to extrapolate data in the future (i.e., after the estimation window), in the context of model predictive control (MPC).

3.1. Experimental example
In this section we introduce the test case for the results that will follow in section 3.2. We will be rather concise, since all details of the experiment can be found in [4, 35]. The test setup consists of a 1 m long aluminum beam clamped on each side to a vertical mount, which is fixed to the ground (Fig. 3). We built a finite element (FE) model and we updated it based on a series of experimental modal analyses (EMAs). For the experiments that we present in this paper, three eigenmodes govern the dynamical behavior of the system, which we report in Fig. 4. The numbering 1, 3, 8 derive from the whole mode set [35]. These dominate the structure response due to the orientation of the external force. In fact, we attached a shaker to the structure such that it applies a force along the z axis, and these modes belong to plane xz (the axis orientation

Figure 2: Graphical interpretation of different approaches to tune the base length of the Fourier dictionary. Legend: MHE time steps (+), original Fourier base length (----) adapted Fourier base length (-----), ad-hoc additional shape functions (———).

Figure 3: Experimental setup.
follows the notation in Fig. 3). Fig. 3 shows also an impedance head, which we employed to validate the force estimation. We imported the model in MATLAB by extracting the FE mass and stiffness matrices and projecting them on modal coordinates [36]. This allowed to build a reduced order state-space model that takes into account only the eigenmodes under examination. Beside the state-space representation we employed camera based measurements, and we tracked the displacements along the z axis of 5 markers that cover uniformly the “BEAM STRIP 2” in Fig. 3. All details including hardware and image processing steps to extract the displacements are documented in [4, 35].

3.2. Autocorrelation
The autocorrelation is one way to detect periodicity in time domain having only a few cycles available [33]. Fig. 5 depicts the basic idea of a CS-MHE scheme that includes the autocorrelation. During the first iteration of the filter the window is not optimal and provokes some spectral leakage. However, the signal contains enough information to detect the base period of the input signal by employing the autocorrelation, and set thus the correct length of the shape functions for the next iteration.

In the remaining part of this section we show an example which makes use of the beam setup described in section 3.1, subjected to a load composed by 2 Fourier components at 64 and 128 Hz, sampled at 512 Hz. The idea is to detect the period of the base frequency of the estimated force and to update the dictionary with this new length \( L_{i+1} \) (where \( i + 1 \) indicates the next iteration) as indicated in eq. (8), where \( \Delta t_{\text{auco}} \) is the period determined by the autocorrelation. \( L_0 \) in eq. (8) can also be replaced by the current length \( L_i \), in case such an update scheme would be more appropriate for a specific application.

\[
L_{i+1} = \text{round}(L_0/\Delta t_{\text{auco}}) \Delta t_{\text{auco}}
\]  

(8)

The graphs in Fig. 6 show the influence of two parameters on the proposed approach. The first parameter is a threshold on the autocorrelation (\( \varepsilon_{\text{auco}} \), i.e., we neglect the peaks of the autocorrelation which are below a certain level) and the second parameter is a threshold on the Fourier components (\( \varepsilon_{\alpha} \), i.e., we do not consider the Fourier components under a certain amplitude [1]). Each column in Fig. 6 corresponds to a different combination of thresholds, indicated in the title of the graphs in the first row. Within each column, from top to bottom the three graphs are:

(i) the number of equivalent time steps \( N_{\text{eq}} \) sampled at \( 1/dt \) that a window generated by eq. 8 would have. \( N_{\text{eq}} \equiv N \) for the first iteration, and it is thus possible to identify the curves by

Figure 4: Mode shapes of the first three beam eigenmodes in plane \( xz \).
Figure 5: Idea of using the autocorrelation to limit spectral leakage as of the second iteration. Amplitude of the Fourier components of the estimated signal (top, the nonzero components are marked by a filled-in circle). Inverse Fourier transform (center). Autocorrelation and peak detection to determine the base length for the next iteration (bottom).

(i) The graphs in the first column refer to a case in which $\varepsilon_{\text{auco}}$ and $\varepsilon_\alpha$ are not well tuned (they are both set quite low, such that the small components are taken into account). This results in a window that does not always follow the correct periodicity (top), a number of nonzero components that does not stabilize to a minimum (middle), and finally a MSE that does not converge to a minimum for all window lengths (bottom).

(ii) For the second column we set a higher threshold on the autocorrelation. We see that the base length of the dictionary is identified correctly, i.e., it stabilizes on three different values from iteration 2 on, according to eq. 8 (top). The number of nonzero components decreases gradually (middle), and the MSE of iteration 2 drops for every window length (bottom).

(iii) The graphs in the third column involve a better tuning where also the threshold on the Fourier components is higher, and the results become as expected: stable detection of the Fourier base length (top), number of nonzero component converging to 5 from iteration 2 on (middle), and MSE dropping starting from iteration 2 (bottom).
Figure 6: Effect of using the autocorrelation for different three combinations of thresholds $\varepsilon_{\text{auco}}$ and $\varepsilon_\alpha$ on the equivalent number of time steps $N_{eq}$, the number of nonzero components $n_\alpha^*$ and on the mean square error (MSE). Every curve refers to a constant $N$, which corresponds to $N_{eq}$ in the first iteration.

Figs. 7–8 show the results of the CS-MHE state/input estimation for the first and second iteration, respectively, in case $N = 34$ time steps. In each figure, the two graphs on the left hand side refer to the state estimation, and depict position and velocity modal participation factors (MPFs), that refer to the three mode shapes in Fig. 4. Moving to the right, the next two graphs show the results of the input estimation, expressed as real part $\Re(\alpha)$ and imaginary part $\Im(\alpha)$ of the Fourier components $\alpha$, and arranged according to the MATLAB® convention for the DFT, i.e., first the DC component, then the half space of positive wave numbers and finally the negative half space. The $\pm 3\sigma$ confidence bands come from the covariance matrix of the constrained optimisation problem [1, 37]. Finally, the two graphs on the right hand side were already proposed in Fig. 5, and are now overlapped with the reference force signal measured by the impedance head (green curves). The top-right graph shows the amplitude of the Fourier components $\text{abs}(\alpha)$, and finally the bottom-right graph is obtained by applying the inverse DFT to the $n_\alpha^*$ nonzero components of $\alpha$, which are marked by a solid circle.

The examples in this section involve a coherent sampling scheme, i.e., there are a finite number of samples in one input period independently of the MHE window length $N$. During our research we tested the autocorrelation also in case of components which are not exact multiples of each other and of the sampling. The main conclusion of this investigation was that the autocorrelation
Figure 7: CS-MHE results at iteration 1 with $N = 34$. Legend (state estimation): 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 thin blue lines that represent the confidence level ($\text{MPF}_n \pm 3\sigma_n$). Legend (input estimation): reference values (—-×); CS-MHE estimation (—-○). The nonzero components are marked by a solid circle.

Figure 8: CS-MHE results at iteration 2 with $N = 34$. Legend: see Fig. 7.
suffers from some limitations. It can detect only the dominant component, and its accuracy depends on the sampling scheme (i.e., it works better if the sampling rate is a multiple of the dominant frequency). The autocorrelation is capable of detecting multiple frequencies, but these should have a comparable amplitude, and this is not common for mechanical systems, where eigenmodes and other dynamical behaviors are scaled with respect to the dominant event. Consequently, the autocorrelation does not fully solve the problem of spectral leakage if several dynamical uncorrelated aspects are relevant. For example, a rotating machine (e.g., a transmission) may have a strong structural behavior that does not match the periodicity given by the rotation speed. In such case an ad-hoc dictionary made of Fourier shapes that match the rotational speed (e.g., detected by an encoder) augmented by some shapes that follow the structural eigenfrequencies may perform better (cf. the fourth graph in Fig. 2). This is interesting but requires extra effort when setting up a filter, since building such dictionary is not trivial. This idea is currently under investigation. If computational performance is not a key element, a further option is to work on a longer window and employ the advanced techniques for periodicity detection proposed in [31, 32]. Finally, we could try to improve the results by adding some extra processing before computing the autocorrelation. Starting from the CS-MHE results, we could run a Kalman filter with a random walk model for the input, and then use the autocorrelation. This would be computationally cheap, all sensitivities are already available, and a backward recursion such as the Rauch-Tung-Striebel smoother is also possible [38, 39].

4. Conclusions
In this paper we summarized the compressive sensing–moving horizon estimator (CS-MHE) for the joint estimation of states and a periodic force input modeled by a Fourier dictionary. In particular, we highlighted the structure of Fourier basis functions in order to illustrate how to remove the link between the MHE estimation window and the dictionary. This goes in the direction of dictionary learning since it allows to tune the basis functions according to a certain periodicity. Such procedure aims at minimizing spectral leakage, which occurs if the MHE window does not match the input periodicity, jeopardizing the estimation results because of poor input modeling and observability issues.

We discussed a few approaches to adapt a Fourier dictionary in response to a certain input periodicity and we listed a few methodologies to determine the periodicity (frequency domain, direct measurement for rotating machinery, time domain), and we focused the attention on the autocorrelation. Specifically, we discussed its applicability with the aid of numerical investigations based on experimental data, and we showed that the method works well under specific settings. The autocorrelation suffers from some limitation in case the sampling scheme does not match the base period of the load and if multiple components with different amplitude have to be detected. The outcome of this paper is twofold. On the one hand we showed how to set the base length of the dictionary independently of the MHE window, with the aim of setting it such that it minimizes spectral leakage. On the other hand we demonstrated that the autocorrelation detects the periodicity only under specific circumstances, leaving space for future research. This may involve further ways to detect periodicity starting from a short signal (possibly including extra steps such as an extended Kalman filter and a Rauch-Tung-Striebel smoother) and may end up in more structured ad-hoc dictionaries and irregular sampling schemes. In the short term we are planning to apply the CS-MHE for the estimation of a quasi-periodic torque in a rotating drive train.

Acknowledgments
This research was partially supported by Flanders Make, the strategic research center for the manufacturing industry, within the MoForM project. The Research Fund KU Leuven is also gratefully acknowledged for its support.
References

[1] Kirchner M, Croes J, Cosco F and Desmet W 2018 Mechanical Systems and Signal Processing 101 237–253
[2] Kirchner M, Croes J, Cosco F, Pluymers B and Desmet W 2016 Proceedings of ISMA2016 including USD2016, Leuven, Belgium ed Sas P, Moens D and van de Walle A pp 2947–2962
[3] Kirchner M, Croes J, Cosco F and Desmet W 2017 Procedia Engineering 199 447–452
[4] Kirchner M, Croes J, Cosco F, Pluymers B and Desmet W 2018 Proceedings of ISMA2018 and USD2018, Leuven, Belgium ed Desmet W, Pluymers B, Moens D and Rottiers W pp 1487–1502
[5] Rao C V and Rawlings J B 2002 AIChE Journal 48 97–109
[6] Haseltine E L and Rawlings J B 2005 Industrial & Engineering Chemistry Research 44 2451–2460
[7] Candes E and Wakin M 2008 Signal Processing Magazine, IEEE 25 21–30
[8] Ginsberg D and Fritzen C P 2014 Proceedings of ISMA2014 including USD2014, Leuven, Belgium ed Sas P, Moens D and Denayer H pp (2014), pp. 2043–2056
[9] Rezayat A, Nassiri V, Pauw B D, Ertveld J, Vanlanduit S and Guillaume P 2016 Mechanical Systems and Signal Processing 70–71 756–768
[10] Qiao B, Mao Z and Chen X 2016 Proceedings of ISMA2016 including USD2016, Leuven, Belgium ed Sas P, Moens D and van de Walle A pp 1685–1696
[11] Qiao B, Zhang X, Gao J and Chen X 2016 Journal of Sound and Vibration 376 72–94
[12] Qiao B, Zhang X, Gao J, Liu R and Chen X 2017 Mechanical Systems and Signal Processing 83 93–115
[13] Qiao B, Zhang X, Wang C, Zhang H and Chen X 2016 Journal of Sound and Vibration 368 71–86
[14] Naets F, Croes J and Desmet W 2015 Computer Methods in Applied Mechanics and Engineering 283 1167–1188
[15] Simon D 2006 Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches (Wiley-Interscience)
[16] Harris F 1978 Proceedings of the IEEE 66 51–83
[17] Kirchner M and Nijman E 2014 Proceedings of ISMA2014 including USD2014, Leuven, Belgium ed Sas P, Moens D and Denayer H pp 1531–1546
[18] Chen S S, Donoho D L and Saunders M A 2001 SIAM Rev. 43 129–159
[19] DeCarlo R A 1989 Linear Systems: A State Variable Approach with Numerical Implementation (Upper Saddle River, NJ, USA: Prentice-Hall, Inc.)
[20] Boyd S and Vandenberghe L 2004 Convex Optimization (New York: Cambridge University Press)
[21] Rao C V, Rawlings J B and Lee J H 2001 Automatica 37 1619–1628
[22] Lopez Negrete de la Fuente R 2011 Nonlinear Programming Sensitivity Based Methods for Constrained State Estimation Ph.D. thesis Carnegie Mellon University (Pittsburgh, PA), Department of Chemical Engineering
[23] Lopez Negrete de la Fuente R and Lorenz B T 2012 Journal of Process Control 22 677–688
[24] Kim S J, Koh K, Lustig M, Boyd S and Gorinevsky D 2007 IEEE Journal of Selected Topics in Signal Processing 1 606–617
[25] Lüöberg J 2004 Proceedings of the CACSD Conference (Taei, Taiwan)
[26] MOSEK ApS 2015 MOSEK Quickstart Guide Version 7.1 (Revision 60)
[27] Duarte-Carvajalino J and Sapiro G 2009 Image Processing, IEEE Transactions on 18 1395–1408
[28] Liao Y, Xiao Q, Ding X and Guo D 2009 International Joint Conference on Computational Sciences and Optimization (CSO 2009) vol 1 pp 831–834
[29] Claeyts T, Vanoost D, Peuteman J, Vandenbosch G A E and Pinsoort D 2015 IEEE Transactions on Electromagnetic Compatibility 57 1329–1337
[30] Claeyts T, Vanoost D, Peuteman J, Vandenbosch G A E and Pinsoort D 2018 IEEE Transactions on Electromagnetic Compatibility 60 202–210
[31] Box G E P and Jenkins G M 1994 Time Series Analysis: Forecasting and Control 3rd ed (Upper Saddle River, NJ, USA: Prentice Hall PTR)
[32] Kirchner M and Nijman E 2016 SpringerBriefs in Applied Sciences and Technology, Automotive NVH Technology ed Fuchs A, Nijman E and Priebsch H (Springer International Publishing) pp 47–62
[33] Kirchner M 2018 Joint state/input estimation in structural dynamics - State/force estimation using compressive sensing within a multistep approach Ph.D. thesis KU Leuven
[34] Heylen W, Lammens S and Sas P 2007 Modal Analysis Theory and Testing (KU Leuven, Department of Mechanical Engineering, PMA Section)
[35] Bock H G, Kostina E and Kostyukova O 2007 SIAM Journal on Matrix Analysis and Applications 29 626–642
[36] Rauch H E, Striebel C T and Tung F 1965 AIAA Journal 3 1445–1450
[37] Chan W L and Hsiao F B 2011 Sensors 11 3738–3764