Online Model Identification for Set-valued State Estimators With Discrete-Time Measurements

João V. Messias
Instituto Superior Técnico
Lisboa, Portugal
Email: jmessias@isr.ist.utl.pt

***Abstract*—Although Kalman Filters provide an optimal solution to the state estimation problem for linear systems, they require knowledge of an accurate description of the system's model. More robust approaches to Kalman Filtering for systems with uncertain models have been developed, such as set-valued state estimators, which identify the set of possible states that the system may be in given its uncertainties. This work presents a practical example of the application of a set-valued state estimator, and also provides insight into the advantages of performing simultaneous model identification for this class of estimators. The results of this robust estimator are then compared to those obtained when using**

I. Introduction

Kalman Filters are a very significant and well-established development in the field of systems and control theory. Given a linear system affected by stochastic noise, the Kalman Filter is able to provide the optimal estimate of the system's state based on gathered noisy observations. However, this class of filters is not robust to unmodeled variations in the system's model. In this case, optimality is lost, and the performance of the filter degrades considerably. Since, in practice, there are situations where it is difficult to estimate the system's model reliably, or it may possess parameters which are impossible to fully specify beforehand, there is a need for estimators which are robust to this type of variations, even if they are suboptimal. Examples of robust filters include Guaranteed Cost State Estimators [1], Set-Valued State Estimators [1], [5], [6] , robust H^{∞} filters [4], [7] and adaptive filters such as the Multiple Model Adaptive Estimator (MMAE) [8]. Of these, the class of set-valued state estimators provide a wide range of functionalities, since they allow robust estimation under dynamic, non-linear uncertainties in the plant and sensor models, by defining a general constraint (Linear Quadratic Constraint) over these uncertainties, based on an uncertainty description introduced in the work of Yakubovich [9]. This class of estimators also allows for on-line model validation, and is suitable for applications involving mixed continuous/discrete measurements (or discrete at different rates) which may possess missing data.

This work studies the application of a set-valued state estimator in a simulated practical scenario, and presents results that demonstrate that it is possible to obtain reliable state estimates in the presence of model uncertainty. Although the authors of this class of estimators explicitly considered the problem of model validation, i.e. to determine if a given model is compatible with the controls applied to the system and the resulting measurements produced by it, its application is restricted to the problem of obtaining a suitable set of possible system models. In practice, it is advantageous to also produce an on-line estimate of the true system model, since this will evidently reduce the amount of uncertainty present in the state estimate. This issue is also explored in this work, by inspecting the behavior of the estimator if the system's uncertainty is re-evaluated online. This is accomplished by resorting to a parallel MMAE, and feeding the resulting estimate of the system's parameters to the set-valued estimator. The resuts with respect to state estimation from these two filters are then compared, noting that while the former uses a stochastic description of system uncertainty, the latter uses a deterministic one. The impact of this fact on filter usability is then discussed.

II. Hybrid Set-Valued State Estimation

This section introduces the basic concepts and techniques behind set-valued state estimators. Consider the linear uncertain system described by equations (1)-(5), where $x(t)$ is the system's state, $u(t)$ is a deterministic control input, $w(t)$, $v_c(t)$, $v_d(t)$ are "uncertainty inputs", $y_c(t)$, $y_d(t_j)$ are the system's measured outputs, $z_c(t)$, $z_d(t_j)$ are the system's "uncertainty outputs", $A(t)$, $B_1(t)$, $B_2(t)$, $K_c(t)$, $G_c(t)$, $C_c(t)$ are bounded piecewise continuous matrix functions, and $K_d(t_j)$, $G_d(t_j)$, $C_d(t_j)$ are matrix sequences.

$$\dot{x}(t) = A(t)x(t) + B_1(t)w(t) + B_2(t)u(t) \tag{1}$$
$$z_c(t) = K_c(t)x(t) + G_c(t)u(t) \tag{2}$$
$$z_d(t_j) = K_d(t_j)x(t_j) + G_d(t_j)u(t_j) \quad \forall j = 1, ..., k \tag{3}$$
$$y_c(t) = C_c(t)x(t) + v_c(t) \tag{4}$$
$$y_d(t_j) = C_d(t_j)x(t_j) + v_d(t_j) \quad \forall j = 1, ..., k \tag{5}$$

Note that the system definition is general enough to include the possibility of both continuous and discrete-time measurements, and could be easily extended to include multiple assynchronous discrete-time sources of information. The uncertainty in this system is assumed to verify that, within a time interval $]0, s]$, for given initial conditions x_0, and for some constant d:

$$(x(0)-x_0)^T P_0^{-1}(x(0)-x_0)+\int_0^s (w(t)^T Q(t)w(t)+ \\ + v_c(t)^T R_c(t)v_c(t))dt+\sum_{t_j\leq s} v_d(t_j)^T R_d(t_j)v_d(t_j) \leq \\ \leq d+\int_0^s ||z_c(t)||^2 dt+\sum_{t_j\leq s}||z_d(t_j)||^2 \quad (6)$$

Where P_0 is an appropriate positive definite matrix, and $Q(t), R_c(t), R_d(t_j)$ are positive semi-definite. The above is known as an Integral Quadratic Constraint. Intuitively, it imposes an upper bound on the uncertainty introduced by $w(t), v_c(t), v_d(t)$, as defined by linear functions $z_c(t), z_d(t_j)$, which converge if the overall system is stable. This allows the uncertainty inputs to be defined as possibly time-varying and non-linear. It also implies that a suitable description for this type of uncertainty is by establishing norm bounds upon each of the model's uncertain terms. It is easy to remark that, assuming that the uncertainty inputs satify the above constraints for every possible state, then the set of possible states at the initial instant is an ellipsoid whose shape is defined by P_0 and d. If this set is bounded for every possible x_0, d, and every history of inputs $u(t)|_0^s$ and measurements $y_c(t)|_0^s, y_d(t)|_0^s$, then the system is said to be "verifiable". If such is the case, the theory behind set-valued state estimator allows for the possible state ellipsoid to be propagated into any given time instant, and it can be shown that the set remains an ellipsoid under these circumstances [1].

The inclusion of both continous and discrete time components in the system's model requires the following auxiliary definition:

$$f(t_j^-)=\lim_{t\to t_j, t<t_j} f(t) \quad (7)$$

With this, and following [1], the design of a set-valued state estimator relates to the following equations:

$$\dot{P}(t)=A(t)P(t)+P(t)A(t)^T+B_1(t)Q(t)^{-1}B_1(t)^T+ \\ +P(t)K_c(t)^T K_c(t)P(t) \quad (8)$$

$$P(t_j)=\left(P(t_j^-)^{-1}+C_d(t_j)^T R_d(t_j)C_d(t_j)\right)^{-1} \\ \forall j=1,...,k \quad (9)$$

These equations are known as a *jump Ricatti equation*, owing to their possible discontinuities at the sampling instants t_j. In [1], it is shown that if equations (8),(9) have a positive definite solution with initial condition $P(0)=P_0$, then the system is verifable. The estimate of the system's state may then be taken as the center of the ellipsoid of possible states, which has the following update equations:

$$\dot{\hat{x}}=\left(A(t)+P(t)K_c(t)^T K_c(t)\right)\hat{x}(t)+ \\ +\left(B_2(t)+P(t)K_c(t)^T G_c(t)\right)u(t) \quad (10)$$

$$\hat{x}(t_j)=\hat{x}(t_j^-)+P(t_j^-)C_d(t_j)^T R_d(t_j)y_d(t_j)- \\ -P(t_j^-)C_d(t_j)^T R_d(t_j)C_d(t_j)\hat{x}(t_j^-) \quad \forall j=1,...,k \quad (11)$$

If the system is verifiable, then another quantity may be defined, which can be seen as a generalized error between the uncertainty outputs of the system and the uncertainty effectively generated when receiving observations $y_d(t_j)$,$y_c(t)$:

$$\rho_s(u,y_d,y_c)=\int_0^s \left\{||K_c(t)\hat{x}(t)+G_c(t)u(t)||^2 \right. \\ \left. -(C_c(t)\hat{x}(t)-y_c(t))^T R_c(t)(C_c(t)\hat{x}(t)-y_c(t))\right\} dt+ \\ +\sum_{t_j\leq s}\left\{||K_d(t_j)\hat{x}(t_j)+G_d(t_j)u(t_j)||^2 \right. \\ \left. -(C_d(t_j)\hat{x}(t_j)-y_d(t_j))^T R_d(t_j)(C_d(t_j)\hat{x}(t_j)-y_d(t_j))\right\} \quad (12)$$

It can then be shown that the set of possible states that the system may be in, at any time instant s, is given by:

$$X(s,x_0,d,u(t)|_0^s,y_c(t)|_0^s,y_d(t)|_0^2)=\{x\in\Re^n : \\ (x-x_0)^T P^{-1}(s)(x-x_0)\leq d+\rho_s(u,y_d,y_c)\} \quad (13)$$

It is also evident that this set is only non-empty if $\rho_s(u,y_d,y_c)\geq -d$, in which case the system is said to be "realizable", which intuitively means that, for the given parameter uncertainties, this set of states is reachable through $u(t)|_0^s$, $y_c(t)|_0^s, y_d(t)|_0^s$, from the initial conditions x_0. The process of model validation refers to using this property to ascertain if a given system model is realizable. In contrast, state estimation in this class of filters requires obtaining the set (13) from the premise that the system is in fact realizable.

It is clear from equations (2),(3), that this type of estimators allows for dynamic re-evaluation of the uncertainty over the system's parameters, by adjusting matrices $K_c(t), K_d(t_j), G_c(t), G_d(t_j)$ accordingly. If these are such that the norm of the uncertainty outputs is minimized, then the set of possible states will also be miminized through (13). However, in [1] the authors do not present a clear methodology to accomplish system identification resorting to the concepts of set-valued estimation. Even so, in order to analyze the behavior of such a system when uncertainty is dynamically re-evaluated, an auxiliary estimator is used for this purpose. The next section introduces an alternative system identification/ state estimation framework that allows for an estimate of the system's parameters to be obtained online.

III. MULTIPLE MODEL ADAPTIVE ESTIMATION

Multiple Model Adaptive Estimation, first proposed by M. Athans in [8], is a versatile technique for state estimation and system indentification for linear systems with unknown parameters. Although not the focus of this work, this algorithm must be briefly described since it will be used as a basis for comparison, and it will serve as an independent system identifier for the state-valued estimator.

In this technique, a set of Kalman Filters is implemented in parallel, and to which an hypothetic value of the unknown parameters is assigned. Given enough filters to cover a region of interest in the unknown parameter space, an *a posteriori* probability measure is then assigned to each of these filters, by inferring over the collected data and respective state estimates, which represents the likelihood of each of the filters being the "correct" one, in the sense that they more accurately describe the true system model. The system's state estimate is then taken as the expected value of this probability mass function, from which the state covariance can also be obtained.

The *a posteriori* probability over the bank of N filters is updated, for each filter i, by:

$$P_i(t_{j+1}) = \frac{\beta_i(t_{j+1})e^{-\frac{1}{2}\mathrm{w}_i(t_{j+1})}}{\sum_{k=0}^{N}\beta_k(t_{j+1})e^{-\frac{1}{2}\mathrm{w}_k(t_{j+1})}P_k(t_j)}P_i(t_j) \quad (14)$$

where, given each filter's residual $r_i(t_{j+1}) \equiv y(t_{j+1}) - \hat{y}_i(t_{j+1}|t_j)$, and the covariance over this residual, $S_i(t_{j+1}) \equiv cov\{r_i(t_{j+1}), r_i(t_{j+1})\}$, the relevant weighting terms are defined as:

$$\beta_i(t_{j+1}) = \frac{1}{2\pi\sqrt{\det S_i(t_{j+1})}} \quad (15)$$

$$\mathrm{w}_i(t_{j+1}) = r_i(t_{j+1})^T S_i(t_{j+1})^{-1} r_i(t_{j+1}) \quad (16)$$

It can then be shown that the p.m.f. over the filter bank will eventually converge until the filter which best matches the true system is assigned probability 1.

IV. IMPLEMENTATION: HUMAN TRANSPORT SYSTEM

The application domain here considered is a simulation of the dynamics of a Human Transport System (HTS), for which the control problem is basically the stabilization of an inverted pendulum. Consider the configuration of the HTS, present in Figure 1, in which its main physical characteristics are presented. A user riding the HTS is represented by a point mass of M_2 at a height $\bar{L}_2$, resulting in an inertia $\bar{I}_2$. The HTS itself possesses mass M_1. A force $u(t)$ is applied by its wheels to keep the system stable. The system's sensors are able to measure its angular position, $\theta(t)$, and velocity, $\dot{\theta}(t)$.

The dynamics of the HTS can be modelled as the nonlinear system:

$$\ddot{\theta} = \frac{1}{m(M_2,\theta)}\left(u - \left(M_2 + \bar{M}_1\right)g\tan(\theta) + M_2\bar{L}_2\sin(\theta)\dot{\theta}^2\right) \quad (17)$$

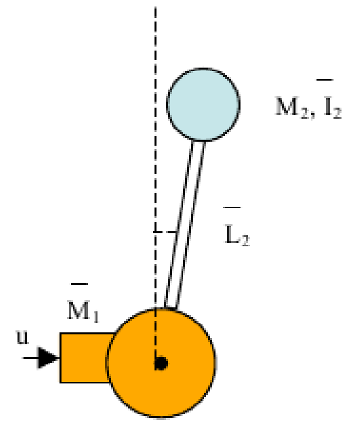

Fig. 1. Representation of the Human Transportation System

where $g = 9.81ms^{-2}$, and

$$m(M_2,\theta) = \left(M_2\bar{L}_2\cos(\theta(t)) - \frac{\left(M_2 + \bar{M}_1\right)\left(M_2\bar{L}_2^2 + \bar{I}_2\right)}{M_2\bar{L}_2\cos(\theta(t))}\right) \quad (18)$$

However, for small angles, these dynamics can be approximated by the linear system:

$$\begin{bmatrix} \dot{\theta}(t) \\ \ddot{\theta}(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \alpha(M_2) & 0 \end{bmatrix}\begin{bmatrix} \theta(t) \\ \dot{\theta}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \beta(M_2) \end{bmatrix}u(t) \quad (19)$$

where $\alpha(M_2)$, $\beta(M_2)$ are nonlinear functions of parameter M_2 (their dependence on M_2 will be ommited). The latter is, therefore, the source of uncertainty in the system's model. It is assumed that the admissible user weights lie in the $M_2 \in [50, 100]kg$ range. As an uncertain system, the HTS can then be approximated by:

$$\dot{x}(t) = \begin{bmatrix} 0 & 1 \\ \alpha & 0 \end{bmatrix}x(t) + \begin{bmatrix} 0 \\ \beta \end{bmatrix}u(t) + \begin{bmatrix} 0 \\ 1 \end{bmatrix}w(t) \quad (20)$$

$$z_c(t) = \begin{bmatrix} k_\alpha(t) & 0 \end{bmatrix}x(t) + k_\beta(t)u(t) \quad (21)$$

$$y_d(t_j) = x(t_j) + v_d(t_j) \quad (22)$$

with the following additional norm constraint for $w(t)$:

$$w(t) = \Delta z_c(t) \qquad ||\Delta|| \leq 1 \quad (23)$$

The observation noise $v_d(t_j)$ is assumed to be white. The initial conditions are assumed to be such that:

$$N = P(0)^{-1} = \begin{bmatrix} 0.2742 & 0 \\ 0 & 0.0027 \end{bmatrix} \quad (24)$$

$$x(0) = \begin{bmatrix} 0.3 & 0 \end{bmatrix}^T \quad (25)$$

$$d = 1 \quad (26)$$

The matrix weighting functions are assumed to be $R_d(t_j) = Q(t) = I$. The sampling period of the sensors is assumed to be $T = 0.01s$.

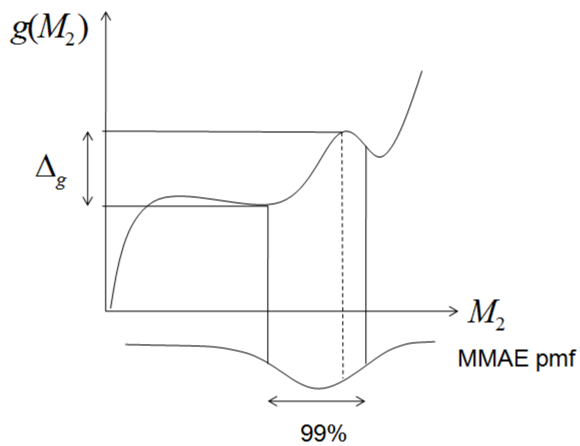

Fig. 2. Estimation of the uncertain parameter's norm bounds from the MMAE user weight estimate.

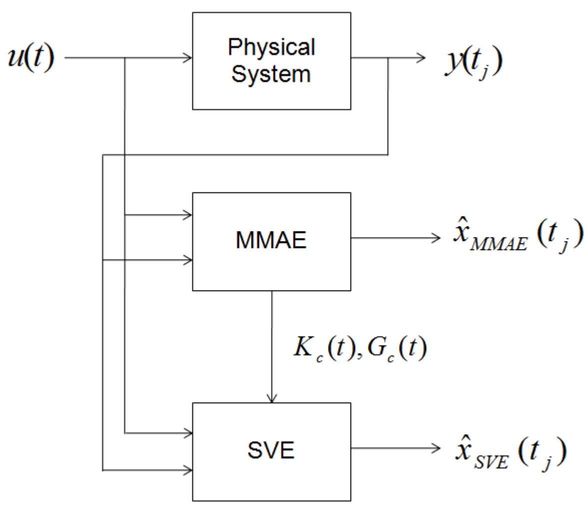

Fig. 3. Connection between the physical system, the SVE, and the auxiliary MMAE.

The absence of continous-time measurements, when noting that the open-loop plant dynamics are unstable within the unknown parameter's region of interest, may lead itself to an unstable estimator if care is not taken into selecting the appropriate values for the weighting matrices. This is intuitively seen since the measurements gathered by the system must adjust the state variables sufficiently to compensate for the unstable predicted motion between two sampling instants.

The $K_c(t)$,$G_c(t)$ coefficients are dynamically estimated by a MMAE by approximating the p.m.f. over its filter bank by a Gaussian distribution, and taking the norm bound of the dependent parameters by establishing a confidence interval over M_2. This process is depicted in figure 2. Figure 3 shows the connection between the "real" system and the Estimators. It should be noted that the MMAE in this particular application serves only as a means of reducing the norm bounds on relevant parameters of the set-valued estimator dynamically. In fact, the purpose of both the set-valued estimator and the MMAE is the same, and, should it be possible, it would be preferrable to estimate these norm bounds within the set-valued estimator itself. Nonetheless, the state estimate returned by the MMAE can also serve as a basis for further comparison.

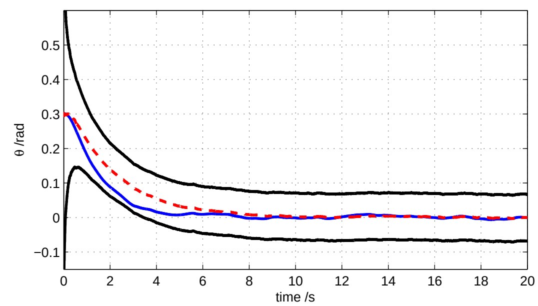

Fig. 4. Evolution of the system's angular position (blue) and respective SVE output (red, dashed). The black bounds denote projections of the limit of the set of possible states. No system identification was performed in this experiment.

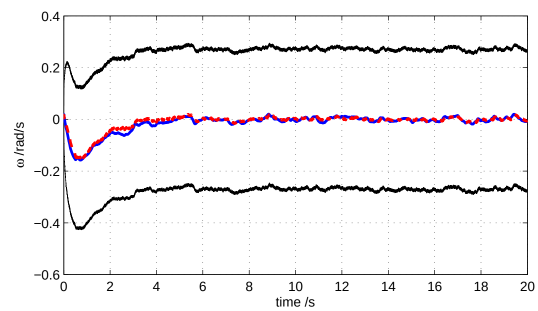

Fig. 5. Evolution of the system's angular velocity (blue) and respective SVE output (red, dashed). The black bounds denote projections of the limit of the set of possible states. No system identification was performed in this experiment.

V. Results

Results were taken from the proposed experimental setup by using MATLAB/Simulink to implement the system. In Figures 4 and 5, the evolution of the state variables is shown, along with the estimate obtained through the SVE and the set bounds, for a situation where there are no external disturbances, and the norm bounds on the uncertainty are held constant. This constitutes a sanity check on the system, and demonstrates that the SVE is able to estimate the state variables correctly, although the possible set bounds are conservative, particularly with respect to the angular velocity of the system.

In Figure 6, a representation of the convergence of the auxiliary MMAE to the correct user weight is shown. From these, it is possible to establish that the uncertainty output's parameters k_α and k_β may also be dynamically estimated by the MMAE. These results are shown in Figures 7 and 8. Note that there is a lower saturation on these estimates, which corresponds to the resolution of the MMAE with respect to the user's weight.

The results of applying these on-line estimates of the uncertain parameters to the SVE are shown in Figures 9 and 10. To make the test more realistic, a sinusoidal disturbance was added to the physical system's output. The independent MMAE state estimate is also shown, which responds considerably better in this situation. This is because the instantaneous observations do not have as much impact in the SVE, since this framework assumes that the sensor model itself may be uncertain. Since the variations induced by the output disturbance still retain the state inside the possible state set bounds, the filter is insensitive to these disturbances.

To better exemplify the response of the system when the its physical parameters are shifted, simulations were run for users of weight $M_2 = 50kg$ and $M_2 = 100kg$. These results are shown in Figures 11 and 12. These show that the estimator is actually more sensitive to the external disturbances in the latter case (these are more pronounced in this case), although the estimated norm bounds enclose the actual state of the system at all times.

The effect of performing online model identification upon the set of possible states at each instant is exemplified in Figure 13. These show that, for a given time instant, the set of possible states is considerably reduced by dynamically reducing the uncertainty norm bounds, as expected. This prompts the suggestion that robust estimation should be accompanied by model identification whenever possible.

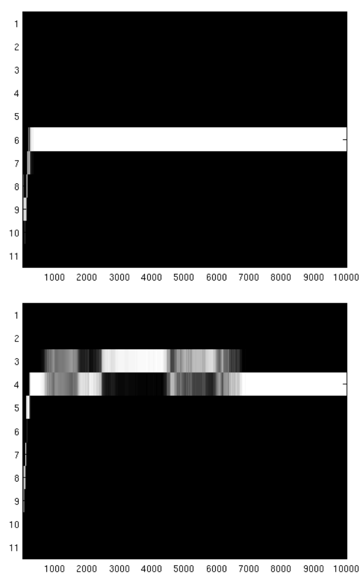

Fig. 6. Convergence of the MMAE p.m.f. Each vertical line in the images represents the p.m.f. over the bank of filters for a given iteration. Lighter colors are associated with higher probability. Top: Real user weight is 75 kg (class 6). Bottom: Real user weight is 65 kg (class 4).

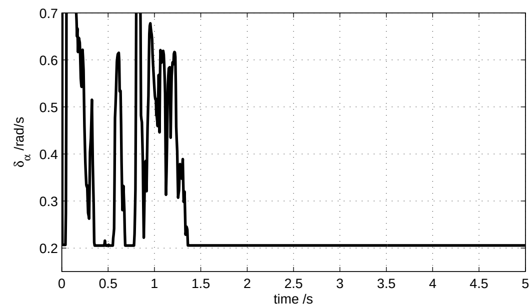

Fig. 7. Evolution of the estimate of the uncertain parameter k_α. After the MMAE converges, the estimate saturates to its lowest admissible value.

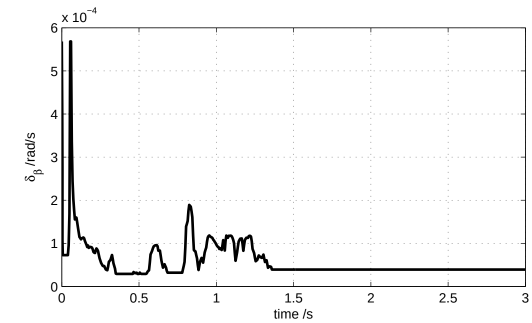

Fig. 8. Evolution of the estimate of the uncertain parameter k_β. After the MMAE converges, the estimate saturates to its lowest admissible value.

VI. Conclusion

In this work, a set-valued state estimator was applied in a practical scenario to demonstrate the capabilites of this framework. The main advantage of this method lies in its deterministic description of uncertainty, since the Integral Quadratic Constraint allows for a large class of uncertainties to be modelled while maintaining the possibility of using a format similar to a Kalman Filter. This allows the set-valued estimator to be more general than when compared to, for example, the MMAE, which would not handle well time-varying uncertain parameters without at least a proper model to describe its evolution.

The robust estimator is not straightforward to implement in practice , since care must be taken into tuning the relevant parameters (such as the uncertainty wheghting terms), that do not have a clear independent meaning. The quality of the estimate produced by the filter will depend largely on the quality of the norm bounds placed over the uncertainty, which can lead to conservative results.

To try to counteract this aspect, the behavior of a set-valued estimator was studied while re-evaluating the system uncertainty online, by performing system identification in an auxiliar estimator. It was verified that the set of possible states is effectively reduced by this process, leading to a filter which is less sensitive to user-set parameters. This also suggests that it may be possible to obtain a suitable method for system identification whithin the SVE framework itself. As it stands, the ability to perform model validation, if a range of possible values for the unknown parameters is already available, is not useful since all models whitin this range are trivially valid. However, a particle-filter based approach, in which the different hypothesis would be different instantiations of the range of the unknown parameters, would likely be useful since

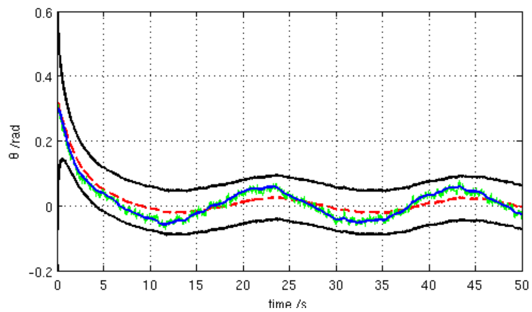

Fig. 9. Evolution of the system's angular position (blue) and respective SVE output (red, dashed). The black bounds denote projections of the limit of the set of possible states. The MMAE estimate is shown in green. System identification was active.

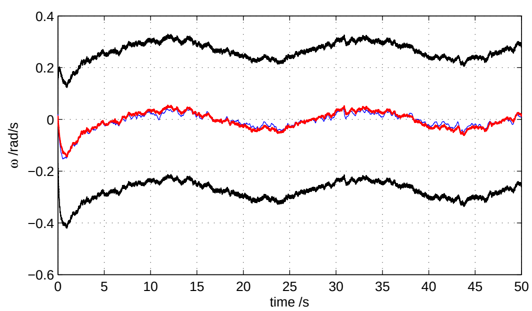

Fig. 10. Evolution of the system's angular velocity (blue) and respective SVE output (red, dashed). The black bounds denote projections of the limit of the set of possible states. System identification was active.

in this case particles could be ruled out by validation. It would be then necessary to define some form of probability measure over the set of possible filters.

References

[1] A. Savkin and I. Petersen, *Robust Kalman Filtering for Signals and Systems with Large Uncertainties*, Birkhäuser, 1999.

[2] P. Bertsekas and I. Rhodes, *Recursive State Estimation for a Set-Membership Description of Uncertainty*, IEEE Transactions on Automatic Control, vol. 16, no. 2, pp. 117-121, 1971.

[3] R. Moheimani, A. Savkin, I. Petersen, *Robust observability for a class of time-varying discrete-time uncertain systems*, Systems & Control Letters, 27, pp.261-266, 1996.

[4] S. Rangan and K. Poolla, *Time-Domain Validation for Sample-Data Uncertainty Models*, In Proceedings of the American Control Conference, pp 1140-1144 Seattle, 1995.

[5] A. Savkin and I. Petersen, *Recursive State Estimation for Unertain Systems with an Integral Quadratic Constraint*, in IEEE Transactions on Automatic Control, vol. 40, no. 6, pp. 1080-1083, 1995.

[6] A. Savkin, I. Petersen and R. Moheimani,*Model validation and state estimation for uncertain continuous-time systems with missing discrete-continuous data*, in Computers and Electrical Engineering, vol. 25, pp. 29-43, 1999.

[7] R. Smith and G. Dullerud, *Continuous-Time Control Model Validation Using Finite Experimental Data*, in IEEE Transactions on Automatic Control, vol. 41, no. 8, pp. 1094-1105, 1996.

[8] M. Athans and C. Chang, *Adaptive Estimation and Parameter Identification using Multiple Model Estimation Algorithm* Technical Note 1976-28, MIT Lincoln Lab, 23 June 1976.

[9] V. A. Yakubovich, *Absolute Stabilization of Nonlinear Systems with a Periodically Nonstationary Linear Part*, in Systems and Control letters, vol 11. no. 3, pp. 221-228, 1988.

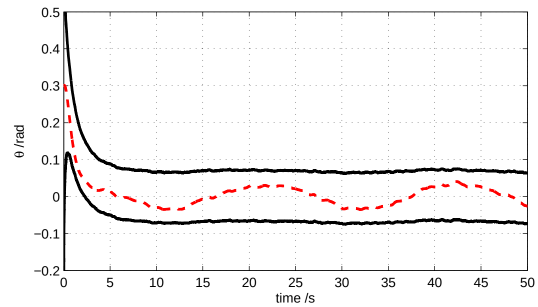

Fig. 11. Actual angular position of the HTS (red) and respective SVE set bounds (black) for a user of weight $M_2 = 50kg$.

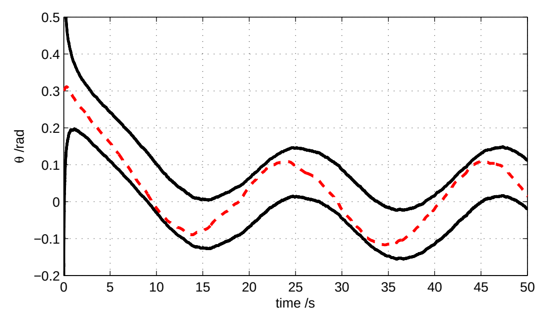

Fig. 12. Actual angular position of the HTS (red) and respective SVE set bounds (black) for a user of weight $M_2 = 100kg$.

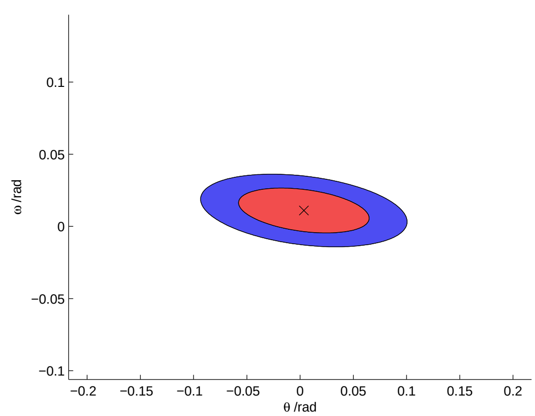

Fig. 13. Representation, for a single time instant $t = 20s$, and for a user weight $M_2 = 75kg$, of the shape of the possible states ellipsoid, when using constant norm bounds (blue) and when performing online system identification (red).