



**HAL**  
open science

## Optimally bounded Interval Kalman filter

Quoc Hung Lu, Soheib Fergani, Carine Jauberthie, Françoise Le Gall

► **To cite this version:**

Quoc Hung Lu, Soheib Fergani, Carine Jauberthie, Françoise Le Gall. Optimally bounded Interval Kalman filter. 2019 IEEE 58th Conference on Decision and Control (CDC), Dec 2019, Nice, France. pp.379-384, 10.1109/CDC40024.2019.9028918 . hal-03300095

**HAL Id: hal-03300095**

**<https://laas.hal.science/hal-03300095>**

Submitted on 8 Jun 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Optimally bounded Interval Kalman filter

Quoc-Hung Lu<sup>1</sup>, Soheib Fergani<sup>1</sup>, Carine Jaubertie<sup>1</sup>, Françoise Le Gall<sup>1</sup>

**Abstract**—This paper is concerned with the optimization of the upper bounds of the interval covariance matrices [1]. This filter is applied to discrete time linear systems subject to mixed uncertainties (combining bounded and stochastic uncertainties), in terms of observations and noises (mainly sensors limitations). It uses interval analysis in order to provide the optimal bound of the state estimation error covariance. Based on that, an optimal state estimation enclosing the set of all possible solutions w.r.t admissible uncertainties is performed. In this article, theorems and lemmas proving the optimality of the proposed solution are provided. Simulations on an example show the efficiency of the developed interval estimation.

## I. INTRODUCTION

State and parameter estimation are topics of utmost importance when dealing with system control. Indeed, obtaining accurate estimations can lead to great improvements in the systems performances. One of the largely used estimation strategy is Kalman filtering. Recently, several extensions of this technique have been presented to deal with the system uncertainties. In [2] and [3], novel designs of optimal robust Kalman filters have been introduced to deal with parameter uncertainties for discrete time-varying systems subject to norm-bounded uncertainties. The main idea of these designs is trying to minimize an upper bound on the estimation error covariance for any acceptable modeling uncertainties. The drawback of these methods is the large conservatism when dealing with too many uncertainties. To overcome this limitation, a solution would be to consider the resulting model matrices, considering the Gaussian noise, as intervals containing all admissible values of parameters. Unfortunately, a singularity problem may appear in interval matrix inversion.

For the last few years, authors have been working on the singularity problem of the interval matrix inversion. While in [1], the solution proposed by using the upper bound of the interval matrix to be inverted is not guaranteed (the solution set may not include all the classical Kalman filter solutions consistent with the bounded uncertainties represented in the system), [4] presented an improved interval Kalman filter (iIKF). It solves the interval matrix inversion problem with the set inversion algorithm SIVIA (Set Inversion Via Interval Analysis) and constraint satisfaction problems (CSP) (see [5]). After that, works focused on reducing both the conservatism and the high computational cost of these algorithms in [6] and [7]. [8] introduced a joint Zonotopic and Gaussian Kalman filter to handle the two paradigms of bounded dis-

turbances and Gaussian noises for discrete-time LTI (Linear Time Invariant) systems. It proposes to perform a multi-objective optimization, minimizing the compromise between the size of the zonotopic part of the bounded disturbances and the covariance of the Gaussian noise distribution.

Motivated by the above observations and the following works in [7], this paper proposes an estimation strategy that provides the optimal upper bound on the state estimation error covariance for all admissible uncertainties. It allows to reduce the size of the set of all possible solutions and improve the computation time of the observations with respect to the classical Kalman filtering structure. The main objective is to estimate the tightest intervals  $[\hat{x}_k]$  that contain the state variables. The upper bound optimization is achieved by an accurate choice of parameters involving the computation of the optimal gain (see section III).

This paper is organized as follows. In Section II, some important definitions, notations and basic concepts needed for the estimation strategy developments are provided. Then, section III presents the main result of the paper which is the Optimal Upper Bound Interval Kalman Filter. Section IV shows the simulations results on a numerical example that prove the efficiency of the proposed estimation strategy. Finally, conclusions and future works are described in section V.

## II. PRELIMINARIES

This section introduces the notations used throughout the paper, including some definitions and properties concerning positive semi-definite matrices and some other basic concepts.

*Definition 1 (Positive semi-definite matrix):* A real matrix  $M \in \mathbb{R}^{n \times n}$  is positive semi-definite if and only if  $M$  satisfies  $z^T M z \geq 0$  for all  $z \in \mathbb{R}^n$ . We denote  $M \succeq 0$ .

All along this article, we only deal with real symmetric positive semi-definite matrices. We denote:

- $S(n) \triangleq \{M \in \mathbb{R}^{n \times n} : M = M^T\}$ , the set of  $n \times n$  symmetric matrices.
- $S_+(n) \triangleq \{M \in S(n) : M \succeq 0\}$ , the set of  $n \times n$  symmetric positive semi-definite matrices.

*Definition 2 (Partial order of real squared matrices):*

Let  $M, N$  be two real squared matrices of the same size. We define an order between  $M$  and  $N$  denoted by  $N \preceq M$  if and only if  $M - N \succeq 0$ .  $M$  is called an upper bound of  $N$ . We also say that  $N$  is dominated by  $M$  or  $M$  dominates  $N$ .

<sup>1</sup> LAAS-CNRS, Université de Toulouse, CNRS, UPS, Toulouse, France. Corresponding author: qhlu@laas.fr

In the case of Hermitian matrices, this order is known as the *Loewner (partial) order* (ref. [9], [10]). Recall that a partial order  $\mathcal{R}$  satisfies the properties: *i) aRa (Reflexivity)*; *ii) If aRb and bRa then a = b (Anti-symmetry)*; *iii) If aRb and bRc then aRc (Transitivity)*.

We also extend this partial order to the notion of bounds for a non empty set  $\Omega$  of real squared matrices. If a real squared matrix  $K$  dominates all matrices contained in  $\Omega$ , then  $K$  is an upper bound of  $\Omega$ , denoted  $\Omega \preceq K$ . In other words,

$$\Omega \preceq K \iff M \preceq K, \forall M \in \Omega.$$

If  $K$  and  $L$  are two upper bounds of  $\Omega$ , then we say that  $K$  is better than  $L$  if and only if the norm of  $K$  is smaller than or equal to the norm of  $L$  depending on the choice of norms in definition 3.

An real interval matrix is a matrix which components are intervals. Let  $[M]$  be an  $n \times n$  real interval matrix. We denote:

- $M \in [M]$  to indicate a punctual matrix  $M$  included in  $[M]$ .
- $S([M]) \triangleq \{M \in [M] : M = M^T\}$ , the set of symmetric matrices included in  $[M]$ .
- $S_+([M]) \triangleq \{M \in S([M]) : M \succeq 0\}$ , the set of symmetric positive semi-definite matrices included in  $[M]$ .
- $BS([M]) \triangleq \{K \in S(n) : S([M]) \preceq K\}$ , the set of symmetric upper bounds of  $S([M])$ .
- $BS_+([M]) \triangleq \{K \in S_+(n) : S_+([M]) \preceq K\}$ , the set of symmetric positive semi-definite upper bounds of  $S_+([M])$ .

In the sequel, we assume that  $S_+([M])$  is non empty. For any  $n \times n$  matrix  $A$ , we use the notations  $\sigma_i(A)$ ,  $\lambda_i(A)$  ( $i = 1, \dots, n$ ) to indicate respectively the singular values and eigenvalues of  $A$  among which  $\sigma_{\max}(A)$  and  $\lambda_{\max}(A)$  are the corresponding maximum values.

*Definition 3:* Let  $A \in \mathbb{R}^{n \times n}$  and  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ . Vector norm and matrix norms are defined as follow. ([10])

a) The Euclidian vector norm:  $\|x\|_2 \triangleq \sqrt{\sum_{i=1}^n x_i^2}$ .

b) The nuclear norm:

$$\|A\|_* \triangleq \sum_{i=1}^n \sigma_i(A) = \sum_{i=1}^n \sqrt{\lambda_i(A^T A)} .$$

c) The operator norm:

$$\|A\| \triangleq \sigma_{\max}(A) = \sqrt{\lambda_{\max}(A^T A)} .$$

d) The Frobenius norm:

$$\begin{aligned} \|A\|_F &\triangleq \sqrt{\sum_{i=1}^n \sigma_i^2(A)} = \sqrt{\sum_{i=1}^n \lambda_i(A^T A)} = \\ &= \sqrt{\text{tr}(A^T A)} = \sqrt{\sum_{i,j=1}^n |A_{ij}|^2}. \end{aligned}$$

*Remark 1:* If  $A \succeq 0$  then  $\sigma_i(A) = \lambda_i(A)$ ,  $\forall i = 1, \dots, n$ , so

$$\begin{aligned} \bullet \|A\|_* &= \sum_i \lambda_i(A) = \text{tr}(A), \\ \bullet \|A\| &= \lambda_{\max}(A), \\ \bullet \|A\|_F &= \sqrt{\sum_{i=1}^n \lambda_i^2(A)} = \sqrt{\text{tr}(A^2)} = \\ &= \sqrt{\sum_{i,j=1}^n |A_{ij}|^2}. \end{aligned}$$

### A. Properties

First of all, we note that  $S(n)$  is a vector (sub-)space (of  $\mathbb{R}^{n \times n}$ ) since  $\alpha M + \beta N \in S(n)$ ,  $\forall M, N \in S(n)$ ,  $\forall \alpha, \beta \in \mathbb{R}$  and  $S_+(n)$  is a convex cone since the above expression is just satisfied for  $\alpha, \beta > 0$ .

*Proposition 1:* Let  $M \in S(n)$  such that  $\lambda_{\max}(M) < \infty$ . Then  $M \preceq \alpha I$  if and only if  $\alpha \geq \lambda_{\max}(M)$ .

*Proposition 2:* Let  $A, B \in S_+(n)$  such that  $A \preceq B$  then

$$\|A\| \leq \|B\| \quad \text{and} \quad \|A\|_* \leq \|B\|_*.$$

For all the following propositions and corollary, let  $[M]$  be an  $n \times n$  real interval bounded symmetric matrix.

*Proposition 3:* The following properties are verified:

- $S([M])$  is compact in the norm vector space  $S(n)$ .
- $S_+([M])$  is a compact subset of  $S([M])$ .
- $\Gamma \triangleq \{\gamma = \|M\| : M \in S_+([M])\}$  is compact in  $\mathbb{R}$ .
- $\sup_{M \in S_+([M])} \{\lambda_{\max}(M)\} < \infty$ .

*Proof:*

- The upper triangular part of matrix  $[M]$  is composed of  $m = \frac{n^2+n}{2}$  intervals  $I_1, \dots, I_m$ . We can construct a continuous function  $f$  from  $I_1 \times \dots \times I_m$  in  $\mathbb{R}^{n \times n}$ . Then, since  $I_1 \times \dots \times I_m$  is compact in  $\mathbb{R}^m$ , the image  $f(I_1 \times \dots \times I_m) = S([M])$  is also compact in  $S(n)$ . The construction of  $f$  is given by:  $f = \psi \circ \phi$  with  $\phi : I_1 \times \dots \times I_m \rightarrow \mathbb{R}^{n \times n}$  verifying  $x = (x_1, \dots, x_m) \mapsto \phi(x) = N =$

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ 0 & x_{n+1} & \cdots & x_{2n-1} \\ \vdots & \cdots & \ddots & \cdots \\ 0 & 0 & \cdots & x_m \end{bmatrix}$$

and  $\psi : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}^{n \times n}$  verifying

$$N \mapsto \psi(N) = N + N^T - \text{diag}(N)$$

are two continuous functions.

- It is only necessary to prove that  $S_+([M])$  is closed in  $S(n)$ , and the result is concluded by the property: *If  $K$  is compact in a topological space  $X$  and if  $F$  is closed in  $X$  with  $F \subseteq K$ , then  $F$  is compact.*

We assume that  $\{M_k\}_k$  is a sequence in  $S_+([M])$  converging to  $M_\infty \in S(n)$  and prove that  $M_\infty \in S_+([M])$ , i.e.  $M_\infty \in [M]$  and  $M_\infty \succeq 0$ .

By assumption,

$$\|M_k - M_\infty\|_F^2 = \sum_{i,j} (M_{k,ij} - M_{\infty,ij})^2 \xrightarrow{k \rightarrow \infty} 0$$

hence

$$(M_{k,ij} - M_{\infty,ij})^2 \xrightarrow{k \rightarrow \infty} 0, \quad \forall i, j = 1, \dots, n.$$

Since each  $M_{k,ij}$  belongs to an (closed) interval  $I_{ij}$  of matrix  $[M]$  then  $M_{\infty,ij} \in I_{ij}$  and  $M_\infty \in [M]$ .

Next, we prove that  $u^T M_k u \xrightarrow{k \rightarrow \infty} u^T M_\infty u, \forall u \in \mathbb{R}^n$ . In fact,

$$\begin{aligned} |u^T M_k u - u^T M_\infty u| &= \left| \sum_{i,j} u_i (M_{k,ij} - M_{\infty,ij}) u_j \right| \\ &\leq \|M_k - M_\infty\|_F \sum_{i,j} |u_i u_j| \end{aligned}$$

and  $\|M_k - M_\infty\|_F \xrightarrow{k \rightarrow \infty} 0$  induce  $u^T M_k u \xrightarrow{k \rightarrow \infty} u^T M_\infty u, \forall u \in \mathbb{R}^n$ . Since for each  $u \in \mathbb{R}^n, u^T M_k u \geq 0$  so it is impossible that  $u^T M_\infty u < 0$ . We conclude that  $u^T M_\infty u \geq 0, \forall u \in \mathbb{R}^n$  or equivalently  $M_\infty \succeq 0$ .

- c) Since the operator norm is a continuous function and  $S_+([M])$  is compact, then  $\Gamma$  is also compact.
- d) The result is induced by extreme value theorem using the compactness of  $\Gamma$ . ■

In the sequel, we denote  $\alpha_* \triangleq \sup_{M \in S_+([M])} \{\lambda_{\max}(M)\}$ .

*Proposition 4:* a)  $S_+([M]) \preceq \alpha I$  if and only if  $\alpha \geq \alpha_*$ .

- b)  $\mathcal{E} \triangleq \{M \in S_+([M]) : \text{diag}(M) = \text{diag}(\text{sup}([M]))\}$  is the non empty set of maximal elements of  $S_+([M])$ .
- c) If  $\mathcal{E}^c \triangleq S_+([M]) \setminus \mathcal{E}$  contains at least two elements  $M, N$  such that  $M_{kl} \neq N_{kl}$  for some tuple  $(k, l) : k \neq l, (k, l = 1, \dots, n)$ , then  $S_+([M])$  has no greatest element.

*Proof:*

- a) Thanks to proposition 1, we get  $\alpha \geq \alpha_*$  if and only if  $M \preceq \alpha I, \forall M \in S_+([M])$ . The proposition 4a) is concluded.
- b) First, let  $M \in S_+([M])$  (which is non empty). If  $M \in \mathcal{E}$  then  $\mathcal{E}$  is non empty. If  $M \notin \mathcal{E}$ , i.e.  $\text{diag}(M) \neq \text{diag}(\text{sup}([M]))$ , we denote  $\hat{M} = M + \Delta$  where  $\Delta = -\text{diag}(M) + \text{diag}(\text{sup}([M]))$ . Then  $\hat{M} \in S_+([M])$  and satisfies  $\text{diag}(\hat{M}) = \text{diag}(\text{sup}([M]))$ . So  $\mathcal{E}$  is non empty since  $\hat{M} \in \mathcal{E}$ . In addition, no matrix  $M \in S_+([M])$  such that  $\text{diag}(M) \neq \text{diag}(\text{sup}([M]))$  is a maximal element of  $S_+([M])$  since such a matrix  $M$  is always dominated by another matrix  $\hat{M} \in \mathcal{E}$ . In other words, any maximal element of  $S_+([M])$  (if it exists) must belong to  $\mathcal{E}$ .  
Next, we prove that any element of  $\mathcal{E}$  is a maximal element of  $S_+([M])$ . In fact, no matrix  $M \in S_+([M]) \setminus \mathcal{E}$  dominates an element  $P \in \mathcal{E}$  since  $\text{tr}(P) \not\leq \text{tr}(M)$  (using proposition 2). Hence, we prove that any two elements in  $\mathcal{E}$  do not dominate each other. Let  $P, Q \in \mathcal{E}$  such that  $P \neq Q$  and  $L = P - Q$ . Then  $L$  satisfies  $L_{ii} = 0$  and

$L_{ij} = P_{ij} - Q_{ij}, \forall i, j = 1, \dots, n$ . To obtain  $Q \preceq P$ , it is necessary:

$$\forall u = (u_1, \dots, u_n) \neq 0, \quad u^T L u = 2 \sum_{i < j} u_i u_j L_{ij} \geq 0.$$

This expression must be satisfied for all following choices of vector  $u$ . Let  $p, q = 1, \dots, n$  and  $p < q$ . By choosing

$$u = \tilde{u} = (\delta_p(1) + \delta_q(1), \dots, \delta_p(n) + \delta_q(n))$$

and

$$u = \hat{u} = (\delta_p(1) - \delta_q(1), \dots, \delta_p(n) - \delta_q(n)),$$

where  $\delta_k(l)$  is the Kronecker delta, it is necessary that:  
 $\tilde{u}^T L \tilde{u} = 2 \sum_{i < j} (\delta_p(i) + \delta_q(i)) (\delta_p(j) + \delta_q(j)) L_{ij} = 2L_{pq} \geq 0,$   
 $\hat{u}^T L \hat{u} = 2 \sum_{i < j} (\delta_p(i) - \delta_q(i)) (\delta_p(j) - \delta_q(j)) L_{ij} = -2L_{pq} \geq 0.$

then,  $L_{pq} = 0, \forall p, q = 1, \dots, n$  and  $p < q$ , which implies a contradiction.

- c) Let  $M, N \in \mathcal{E}^c$  such that  $M_{kl} \neq N_{kl}$  for some tuple  $(k, l) : k \neq l, (k, l = 1, \dots, n)$ . Let  $P = M - \text{diag}(M) + \text{diag}(\text{sup}([M]))$  and  $Q = N - \text{diag}(N) + \text{diag}(\text{sup}([M]))$ . Then  $P$  and  $Q$  are such that  $P, Q \in S_+([M]), P \neq Q$  and  $\text{diag}(P) = \text{diag}(Q) = \text{diag}(\text{sup}([M]))$ . Hence  $P$  and  $Q$  belong to  $\mathcal{E}$  and are two different maximal elements of  $S_+([M])$ . This implies that  $S_+([M])$  does not have the greatest element. ■

*Corollary 1:* There exists a matrix  $N^* \in \mathcal{E}$  such that  $\lambda_{\max}(N^*) = \alpha_*$ .

*Proposition 5:* Let  $K \in BS_+([M])$ . Then:

- a)  $\alpha_* I$  is the optimal upper bound of  $S_+([M])$  in the set  $BS_+([M])$  in the sense of operator norm minimization.
- b)  $\alpha_* I$  is the optimal upper bound of  $S_+([M])$  in the set of upper bounds  $K \in BS_+([M])$  such that

$$\frac{\lambda_1(K) + \dots + \lambda_n(K)}{n} \geq \alpha_*$$

in the sense of nuclear norm minimization.

In the following, the notation  $\text{mid}([M])$  stands for the matrix of center points.

*Proposition 6:* Let  $Max$  be a matrix determined by

$$Max_{ij} = \begin{cases} \text{sup}([M])_{ij} & , \text{ if } \text{mid}([M])_{ij} \geq 0 \\ \text{inf}([M])_{ij} & , \text{ otherwise} \end{cases}$$

then  $\sup_{M \in S_+([M])} \{\|M\|_F\} \leq \|Max\|_F$  and  $\alpha_* \leq \|Max\|_F$ . In addition, if  $Max \succeq 0$ , then

$$\lambda_{\max}(Max) \leq \alpha_* \leq \|Max\|_F.$$

### III. MAIN RESULTS

We have proved that a non empty set  $S_+([M])$  can be dominated by an upper bound of the form  $\alpha I$  ( $\alpha > 0$ ). Furthermore, we pointed out that the particular upper bound  $\alpha_* I$  is the optimal bound for  $S_+([M])$  among other bounds in  $BS_+([M])$  according to the operator norm. And if we use the nuclear norm as a criterion,  $\alpha_* I$  is also the optimal bound in a rather large subset of  $BS_+([M])$ : the set of matrices  $K \in BS_+([M])$  such that

$$\frac{\lambda_1(K) + \dots + \lambda_n(K)}{n} \geq \alpha_*.$$

The use of this particular upper bound ensures on one hand a solution to the singularity problem of the interval matrix inversion as well as the conservatism of the algorithms, and on the other hand an appreciable reduction of the computation time. This fact is verified in a numerical simulation in which the use of this kind of upper bound reduces by over 40% the computation time and ensures more consistent estimate intervals. In addition, by using an appropriate choice of coefficients ( $\beta_k$  and  $\sigma_k$ ) the traces of estimated covariance error bounds are well controlled. Although we did not find the exact value of  $\alpha_*$  but a bound for it, the numerical simulation showed that we can rather use its upper bound ( $\|Max\|_F$ ) to obtain the desired results presented in section IV.

From a statistical point of view, we used this particular upper bound as a covariance matrix of a multi-dimensional (Gaussian) noise which components are mutually independent. In other words, in our problem, all covariance matrices included in an interval  $[Q]$  can be represented by the one of a multi-dimensional variable which components are mutually independent. These results are used to develop our new Optimal Upper Bound Interval Kalman Filter.

#### A. Optimal Upper Bound Interval Kalman Filter

Consider the linear discrete time dynamic system represented by the states and measures equations

$$\begin{cases} x_k = A_k x_{k-1} + B_k u_k + w_k \\ y_k = C_k x_k + v_k \end{cases} \quad (1)$$

where the notations are usual for the classic Kalman filter:  $x_k \in \mathbb{R}^{n_x}$  state variables,  $y_k \in \mathbb{R}^{n_y}$  measures,  $u_k \in \mathbb{R}^{n_u}$  inputs,  $w_k \in \mathbb{R}^{n_x}$  state noises,  $v_k \in \mathbb{R}^{n_y}$  measure noises. Matrices  $A_k, B_k, C_k$  are unknown, deterministic and included in bounded interval matrices  $[A], [B], [C]$  respectively.  $w_k, v_k$  are centered Gaussian with covariance matrices  $Q_k$  and  $R_k$  included respectively in bounded interval matrices  $[Q]$  and  $[R]$ . The initial state  $x_0$  is also Gaussian. In addition,  $x_0, \{w_1 : w_k\}$  and  $\{v_1 : v_k\}$  are assumed to be mutually independent.

Our aim is to estimate intervals  $[\hat{x}_k]$  which contain state variables  $x_k$ . Moreover, each value  $\hat{x}_k \in [\hat{x}_k]$  can be considered as a point estimate of  $x_k$ . By applying the Kalman classic filter calculations, we obtain,  $\forall A_k \in [A], \forall C_k \in$

$[C], \forall Q_k \in [Q], \forall R_k \in [R], \forall \hat{x}_k \in [\hat{x}_k]$  and  $\forall k \geq 1$ :

$$\begin{aligned} P_{k|k} &= \mathbb{E}[(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T] \\ &= (I - K_k C_k) \underbrace{(A_k P_{k-1|k-1} A_k^T + Q_k)}_{P_{k|k-1}} (I - K_k C_k)^T \\ &\quad + K_k R_k K_k^T. \end{aligned} \quad (2)$$

Let  $S_+([P_{k|k-1}]) \preceq \alpha_k I$  and  $S_+([R]) \preceq \gamma I$ , then

$$P_{k|k} \preceq \alpha_k (I - K_k C_k) (I - K_k C_k)^T + \gamma K_k K_k^T.$$

By applying the same strategy as the one in [7], we obtain,  $\forall A_k \in [A], \forall C_k \in [C], \forall Q_k \in [Q], \forall R_k \in [R], \forall \hat{x}_k \in [\hat{x}_k]$  and  $\forall k \geq 1$ :

$$\begin{aligned} P_{k|k} &\preceq \alpha_k \left( 1 + \sum_{i=1}^{n_y} \sum_{j=1}^{n_x} \beta_{ij,k}^{-1} T_{ij} \right) (I - K_k m) (I - K_k m)^T \\ &\quad + \sum_{i=1}^{n_y} \sum_{j=1}^{n_x} T_{ij} \left( \beta_{ij,k} + \sum_{u=1}^{n_y} \sum_{v=1}^{n_x} T_{uv} \sigma_{ijuv,k} \right) K_k r_{ij} r_{ij}^T K_k^T \\ &\quad + \gamma K_k K_k^T, \end{aligned}$$

for any  $\beta_{ij,k} > 0, \sigma_{ijuv,k} > 0$  and where  $m = \text{mid}([C])$ ,  $r_{ij} = \text{rad}([C]_{ij})$  and  $T_{ij} = 1$  if  $\text{rad}([C]_{ij}) > 0$  and null otherwise. Note that the notations  $\text{mid}([C])$  and  $\text{rad}([C])$  stand for the matrix of center points and the matrix of radius of  $[C]$  respectively.

We choose  $\beta_{ij,k} = \beta_k > 0$  and  $\sigma_{ijuv,k} = \sigma_k > 0$ , for all  $i, j, u, v$ . Then

$$\begin{aligned} P_{k|k} &\preceq \alpha_k (1 + n_0 \beta_k^{-1}) (I - K_k m) (I - K_k m)^T \\ &\quad + (\beta_k + n_0 \sigma_k) K_k \left( \sum_{i=1}^{n_y} \sum_{j=1}^{n_x} r_{ij} r_{ij}^T \right) K_k^T \\ &\quad + \gamma K_k K_k^T, \end{aligned} \quad (3)$$

where  $n_0$  is the number of non null radius of  $[C]$ . The right term of (3) is denoted by  $\mathcal{P}_{k|k}$ . The optimal gain  $K_k$  which minimizes the trace of  $\mathcal{P}_{k|k}$  is

$$K_k^* = m^T (m m^T + u_k D + v_k I)^{-1}$$

where  $D = \sum_{i=1}^{n_y} \sum_{j=1}^{n_x} r_{ij} r_{ij}^T$ ,  $u_k = \frac{\beta_k + n_0 \sigma_k}{1 + n_0 \beta_k^{-1}}$  and  $v_k = \frac{\gamma}{\alpha_k (1 + n_0 \beta_k^{-1})}$ . By choosing  $K_k = K_k^*$  we get

$$\mathcal{P}_{k|k} = \alpha_k (1 + n_0 \beta_k^{-1}) (I - K_k^* m).$$

The trace of  $\mathcal{P}_{k|k}$  is optimal according to the choice  $K_k = K_k^*$ , but we can do better by well choosing the coefficients  $\beta_k$  and  $\sigma_k$ . We have

$$\begin{aligned} \text{tr}(\mathcal{P}_{k|k}) &= \alpha_k \left( 1 + \frac{n_0}{\beta_k} \right) \text{tr} \left[ I - m^T (m m^T + u_k D + v_k I)^{-1} m \right] \\ &= \alpha_k \left( 1 + \frac{n_0}{\beta_k} \right) \left[ \text{tr}(I) - \text{tr} \left( m m^T (m m^T + u_k D + v_k I)^{-1} \right) \right] \\ &= \alpha_k (\beta_k + n_0 \sigma_k) \text{tr} \left( D (m m^T + u_k D + v_k I)^{-1} \right) \\ &\quad + \gamma \text{tr} \left( (m m^T + u_k D + v_k I)^{-1} \right) \end{aligned}$$

We observe that if  $\beta_k$  and  $\sigma_k$  converge to 0, then  $\text{tr}(\mathcal{P}_{k|k})$  converges to  $\gamma \text{tr} \left( (m m^T)^{-1} \right)$ . The use of upper bounds of  $P_{k|k-1}$  and  $[R]$  with the forms  $\alpha I$  ( $\alpha > 0$ ) decreases

the computation time of the algorithm (with a rate more than 40% according to a numerical result). In addition, by choosing  $\beta_k$  and  $\sigma_k$  sufficient small, we can control the trace of  $\mathcal{P}_{k|k}$ .

#### IV. NUMERICAL EXAMPLE

We simulate an example described by equation (1) with no input, where:

$$[A] = \begin{pmatrix} [2.45, 2.72] & [-1.41, -1.28] & [0.26, 0.28] \\ [6.32, 6.98] & [-3.56, -3.22] & [2.45, 2.72] \\ [-0.79, -0.72] & [0.3, 0.34] & [0.1, 0.11] \end{pmatrix},$$

$$[C] = \begin{pmatrix} [-8.16, -7.84] & [-4.08, -3.92] & [1.96, 2.04] \\ [-2.04, -1.96] & [1.96, 2.04] & [5.88, 6.12] \\ [-0.41, -0.39] & [15.68, 16.32] & [6.86, 7.14] \end{pmatrix},$$

$$[Q] = [R] = \begin{pmatrix} [8, 12] & [-6, -4] & [3.2, 4.8] \\ [-6, -4] & [8, 12] & [1.6, 2.4] \\ [3.2, 4.8] & [1.6, 2.4] & [8, 12] \end{pmatrix}.$$

The initial state is  $x_0 = (5, -2, 6)^T$  and the model starts at  $[\hat{x}_0] = ([-2, 2], [-2, 2], [-2, 2])^T$ . The initial error covariance bound is  $\mathcal{P}_{0|0} = 10I$ .

First, we simulate with  $N = 10000$  steps for state variables  $x_k$ , measures  $y_k$  and covariance matrices  $P_{k|k}$ . More precisely, at each step  $k$ , from  $[A], [C], [Q], [R]$ , we generate respectively matrices  $A_k, C_k, Q_k, R_k$  such that  $Q_k$  and  $R_k$  are symmetric positive semi-definite. Then  $w_k, v_k$  are simulated such that  $w_k \sim \mathcal{N}(0, Q_k)$  and  $v_k \sim \mathcal{N}(0, R_k)$ . From these informations, we can calculate  $x_k, y_k$  and  $P_{k|k}$ . In a second phase, we run our algorithms, named OUBIKF (Optimal Upper Bound Interval Kalman Filter), together with the one of [7], named UBIKF, for  $N$  steps to obtain  $[\hat{x}_k^{opt}]$ ,  $\mathcal{P}_{k|k}^{opt}$  and  $[\hat{x}_k]$ ,  $\mathcal{P}_{k|k}$  respectively. The following results are adapted to the choice of upper bounds  $\alpha I$  with  $\alpha = \|Max\|_F$  and of coefficients  $\beta_k = \frac{1}{2.n_0.10^3}$  and  $\sigma_k = \frac{1}{n_0.10^3}$ .

#### Simulations results

The computation time of OUBIKF is reduced more than 40% w.r.t the one of UBIKF (see Table I).

	RMSE			Time
	$x_1$	$x_2$	$x_3$	
UBIKF	413.41	448.83	343.82	51.375 s
OUBIKF	416.95	451.48	346.51	28.719 s

TABLE I: Results of the OUBIKF versus UBIKF

The traces of bounds  $\mathcal{P}_{k|k}^{opt}$  decrease rapidly and have a convergence tendency while the traces of  $\mathcal{P}_{k|k}$  increase (although bounded) (Fig. 1, Fig. 2).

In addition, we have  $\text{tr}(P_{k|k}) \leq \text{tr}(\mathcal{P}_{k|k}^{opt}) \leq \text{tr}(\mathcal{P}_{k|k})$  for all  $k \geq 1$ . This fact is verified experimentally for all  $k \geq 1$  and a summary is presented in Table II. Furthermore, all estimate intervals  $[\hat{x}_k^{opt}]$  are contained in the corresponding estimate intervals  $[\hat{x}_k]$ .

The next result concerns the confidence intervals defined by

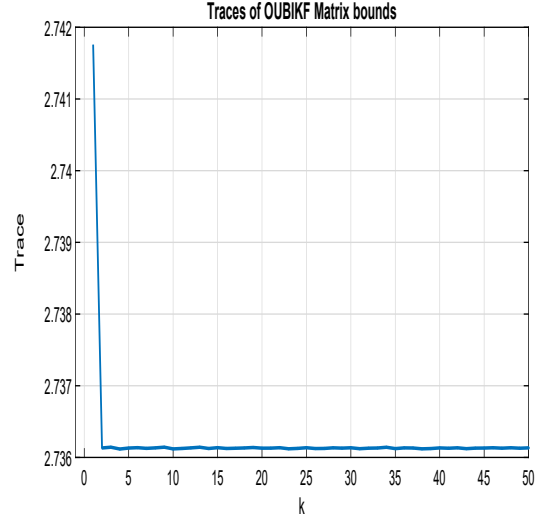


Fig. 1: Behavior of Traces of OUBIKF Matrix bounds

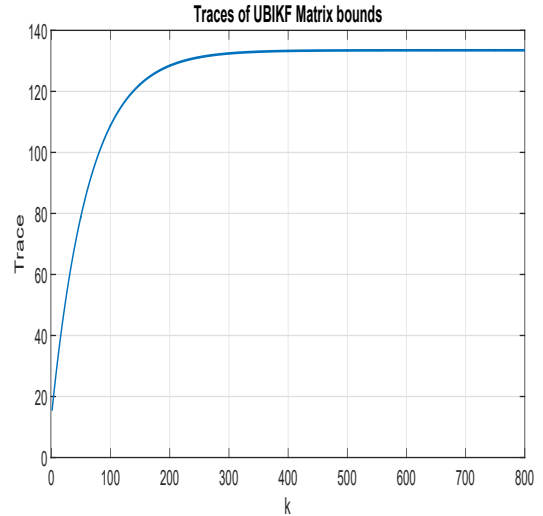


Fig. 2: Behavior of Traces of UBIKF Matrix bounds

$$CI_{k,i}^{opt} = \left[ \inf([\hat{x}_{k,i}^{opt}]) - r\sqrt{\mathcal{P}_{k,ii}^{opt}}, \sup([\hat{x}_{k,i}^{opt}]) + r\sqrt{\mathcal{P}_{k,ii}^{opt}} \right],$$

$$CI_{k,i} = \left[ \inf([\hat{x}_{k,i}]) - r\sqrt{\mathcal{P}_{k,ii}}, \sup([\hat{x}_{k,i}]) + r\sqrt{\mathcal{P}_{k,ii}} \right]$$

for  $i = 1, \dots, n_x$  and  $r = 1, 2, 3$  corresponding to 68%, 95%, 99.7% confidence interval (the 3-sigma rule). According to the simulation, the 68% confidence intervals contain all corresponding state variables  $x_k$  and, more over,  $CI_{k,i}^{opt} \subseteq CI_{k,i}, \forall k \geq 1, \forall i = 1, \dots, 3$  (Fig. 3).

So the  $O(\%)$  which is the percentage of confidence intervals containing corresponding state variables are both 100% for two algorithms, however the  $CI_{k,i}^{opt}$ 's are more consistent.

We note that the 68% confidence intervals do not always contain the corresponding state variables with  $O(\%) = 100\%$  for different simulations (with the same input), but often reach over 99.9%. And until now, none of our simulations for

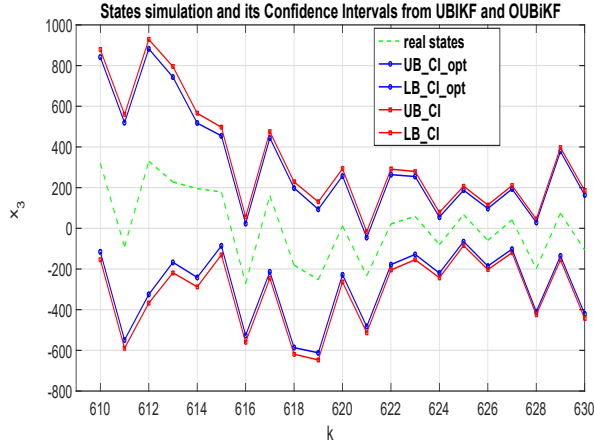


Fig. 3: State variable and its 68% Confidence Intervals

this example with 99.7% confidence interval gets the  $O(\%)$  under 100%.

Concerning the meaning of the confidence interval definition presented above, we can think in the manner that  $CI_{k,i} = \bigcup_{\hat{x}_{k,i} \in [\hat{x}_{k,i}]} CI_{k,i}(\hat{x}_{k,i})$  where  $CI_{k,i}(\hat{x}_{k,i}) = [\hat{x}_{k,i} - r\sqrt{\mathcal{P}_{k,ii}}, \hat{x}_{k,i} + r\sqrt{\mathcal{P}_{k,ii}}]$ . A state  $x_k$  belonging to  $CI_{k,i}$  means that  $x_k \in CI_{k,i}(\hat{x}_{k,i})$  for some  $\hat{x}_{k,i} \in [\hat{x}_{k,i}]$ . It is the same thing for  $CI_{k,i}^{opt}$ .

We also deal with a criterion called Root Mean Squared Error (RMSE) to compare the performance of the two algorithms. The RMSE is defined by

$$RMSE_i = \sqrt{\frac{1}{N} \sum_{k=1}^N (x_{k,i} - \text{mid}([\hat{x}_{k,i}]))^2}$$

for  $i = 1, \dots, n_x$ . And the result is that the RMSE for OUBIKF is slightly increasing w.r.t the one for UBIKF (Table I). We use this criterion since it has been used before to compare different algorithms, e.g. in [7]. But a critical point of view can be pointed out. The distance between the state  $x_{k,i}$  and the center point of the corresponding estimate interval is used. This fact dismisses the issue of the estimate interval width. Naturally, two estimate intervals with the same center point have the same RMSE. In other words, this index just stand for the concentration of state  $x_k$  with respect to the corresponding estimate interval center point. We propose another distance to improve the meaning of this criterion. We will use the Hausdorff distance defined as follow:

$$d_H(X, Y) = \max \left\{ \sup_{x \in X} \inf_{y \in Y} d(x, y), \sup_{y \in Y} \inf_{x \in X} d(x, y) \right\}$$

for any two non empty sets  $X, Y$  in the metric space  $(\Omega, d)$ . In our case, we have

$$d_H(x_{k,i}, [\hat{x}_{k,i}]) = \max \{ |x_{k,i} - \inf([\hat{x}_{k,i}])|, |x_{k,i} - \sup([\hat{x}_{k,i}])| \}$$

and the new RMSE is defined by

$$\widehat{RMSE}_i = \sqrt{\frac{1}{N} \sum_{k=1}^N d_H(x_{k,i}, [\hat{x}_{k,i}])^2}$$

The result for  $\widehat{RMSE}$  in Table III shows that the estimate intervals  $[\hat{x}_{k,i}^{opt}]$  are more relevant by their tightness.

Trace	Min	Mean	Max	Width*
$\text{tr}(P_{k k})$	1.0592	1.3399	1.6418	0.5826
$\text{tr}(P_{k k}^{opt})$	2.7361	2.7361	2.7418	0.0057
$\text{tr}(P_{b_{k k}})$	15.353	132.72	133.47	118.117

\*Width = Max - Min

TABLE II: Traces of Covariance matrices and Traces of its Upper bounds

	$\widehat{RMSE}$		
	$x_1$	$x_2$	$x_3$
UBIKF	5047.2	4159.1	4213.6
OUBIKF	4708.5	3847.7	3934.2

TABLE III: The  $\widehat{RMSE}$

## V. CONCLUSION

This paper proposes a novel result to find an optimal upper bound of set  $S_+([M])$  that ensures a solution to the singularity problem of the interval matrix inversion as well as the conservatism of the algorithms. Based on that result, an Optimal Upper Bound Interval Kalman Filter (OUBIKF) is developed to perform an optimal state estimation that encloses the set of all possible solutions w.r.t admissible uncertainties. Simulations results shows the efficiency and accuracy of the proposed strategy.

## REFERENCES

- [1] G. Chen, J. Wang, and S. Shieh, "Interval Kalman filtering," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 33, no. 1, pp. 250–259, Jan 1997.
- [2] D. Zhe and Y. Zheng, "Finite-horizon robust Kalman filtering for uncertain discrete time-varying systems with uncertain-covariance white noises," *IEEE Signal Processing Letters*, vol. 13, no. 8, pp. 493–496, Aug 2006.
- [3] S. Mohamed and S. Nahavandi, "Robust finite-horizon Kalman filtering for uncertain discrete-time systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1548–1552, June 2012.
- [4] J. Xiong, C. Jauberthie, L. Travé-Massuyès, and F. Le Gall, "Fault detection using interval Kalman filtering enhanced by constraint propagation," in *Proceedings of the 52nd IEEE Annual Conference on Decision and Control (CDC)*, Florence, Italy, Dec 2013, pp. 490–495.
- [5] L. Jaulin, M. Kieffer, O. Didrit, and E. Walter, *Applied Interval Analysis, with Examples in Parameter and State Estimation, Robust Control and Robotics*. London: Springer-Verlag, 2001.
- [6] T. A. Tran, F. Le Gall, C. Jauberthie, and L. Travé-Massuyès, "Two stochastic filters and their interval extensions," in *Proceedings of the 4th IFAC International Conference on Intelligent Control and Automation Sciences*, Reims, France, June 2016.
- [7] T. A. Tran, C. Jauberthie, F. Le Gall, and L. Travé-Massuyès, "Interval kalman filter enhanced by positive definite upper bounds," in *Proceedings of 20th IFAC World Congress 2017*.
- [8] C. Combastel, "Merging Kalman filtering and zonotopic state bounding for robust fault detection under noisy environment," in *Proceedings of the 9th IFAC Symposium on Fault Detection, Supervision and Safety of Technical Processes*. Paris, France, sep 2015.
- [9] F. Pukelsheim, *Optimal design of Experiments*, ser. SIAM's Classics in Applied Mathematics. New York: M. Dekker, 2006.
- [10] X. Zhan, *Matrix Inequalities*. Springer-Verlag Berlin Heidelberg, 2002.