State-bounding estimation for nonlinear models with multiple measurements

Y. Becis-Aubry and N. Ramdani

Abstract—A hierarchical state bounding estimation method is presented for nonlinear dynamic systems where different sensors offer several measurements of the same state vector, each of which is subject to unknown but bounded disturbances and is equipped with a local processor.

For each sampling time, the proposed algorithm proceeds in two stages.

At the prediction stage, an approximating outer-bounding ellipsoid is computed for the reachable set of the nonlinear function of the state vector.

At the correction stage, the algorithm works at two levels: Each local processor computes the state estimate and its outer-bounding ellipsoid according to the local measurements given by the corresponding sensor. These ellipsoids are transmitted simultaneously from all local processors to the fusion center which synthesizes them to compute the global state bounding ellipsoid. Then it feeds these data back to all the local processors. This feedback allows the local processors to adjust their results by taking into account the measurements of all the other sensors.

I. INTRODUCTION

State estimation in a multisensor and multiprocessor framework has been extensively studied during the last three decades under the assumption of white and Gaussian measurement and/or process noises. The fused Kalman filtering presented in [1], [2], [3] and other references within are good examples for that.

In this paper, we suppose that the system and measurement noises have not any known statistical property, they are only supposed to belong to ellipsoidal sets.

To our knowledge, only few approaches have been developed in this field, i.e., when the noises are rather bounded than stochastic, especially when the sets characterizing these bounds are ellipsoids and when the state dynamics are nonlinear.

Recently, an LMI-based state estimation method for linear systems subject to bounded disturbances and model uncertainties has been presented in [4]. In [5], a set-membership estimation algorithm for nonlinear systems with ellipsoidal noise and state characterizations using Takagi-Sugeno fuzzy modeling approach and the S-procedure technique is proposed. In [6], a recursive error-constrained filtering based on the semi-definite program method is developed for nonlinear time-delay systems.

The present work proposes a state estimation approach for models corrupted by bounded disturbances with ellipsoidal noise and state characterizations and complements the one presented in [7] which dealt with multisensor estimation fusion for systems with nonlinear dynamics and with linear-in-state potentially faulty measurements. Here, the multisensor measurements are supposed to be fault-free (cf. Section IV) and a detailed, easy and ready to apply method is presented for the computation of upper bounds, in the sense of positive semi-definiteness, of the Jacobian matrix of the state evolution function which leads to the outer-bounding ellipsoid for the reachable set of this function (cf. Section III). This part was lacking in [7].

Notations and preliminaries:

i. The symbol := (resp. =:) means that the Left Hand Side (resp. RHS) is defined to be equal to the Right Hand Side (resp. LHS).

ii. Let $A \in \mathbb{R}^{n \times m}$, $A := [a_{ij}]$, where $a_{ij}$ is the element of $i^{\text{th}}$ row and $j^{\text{th}}$ column. The elementwise inequality $A < B$ (resp. $A \leq B$) means that $a_{ij} < b_{ij}$ (resp. $a_{ij} \leq b_{ij}$), for all $i \in \{1, \ldots, n\}$ and all $j \in \{1, \ldots, m\}$.

iii. $0$ denotes a zero vector of appropriate dimension and $I_n$ the identity matrix of dimension $n$.

iv. Let $N \in \mathbb{R}^{m \times n}$. $\sigma_i(N)$ is its $i^{\text{th}}$ largest singular value: $\sigma_i(N) = \sigma_{\max}(N)$ and $\sigma_{\min(m,n)}(N) = \sigma_{\min}(N)$.

v. Let $M$ be a square matrix. $\lambda_i(M)$ denotes its $i^{\text{th}}$ (largest if $M$ is symmetric) eigenvalue, $\text{tr}(M) = \sum\lambda_i(M)$ is its trace and $\det(M) = \prod\lambda_i(M)$ is its determinant.

vi. $\|\cdot\|$ is the 2-norm: for any vector $x$, $\|x\|^2 := x^T x$ and for a matrix $A$, $\|A\| := \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} = \sigma_{\max}(A)$.

vii. A Symmetric matrix $M \in \mathbb{R}^{n \times n}$ is Positive Definite, denoted by SPD, (resp. Positive Semi-Definite, denoted by SPSD) if and only if $\forall x \in \mathbb{R}^n \setminus \{0\}$, $x^T M x > 0$ (resp. $x^T M x \geq 0$). This condition is met if and only if all its eigenvalues are real (because of its symmetry) and positive (resp. non-negative).

viii. $E(c, P) := \{ x \in \mathbb{R}^n | (x - c)^T P^{-1} (x - c) \leq 1 \}$ is an ellipsoid in $\mathbb{R}^n$ ($s \in \mathbb{R}^n$), where $c \in \mathbb{R}^n$ is its center and $P \in \mathbb{R}^{n \times n}$ is a SPD matrix that defines its shape, size and orientation in the $\mathbb{R}^n$ space. If the $P$ is only SPSD, then the ellipsoid is degenerate.

ix. $\oplus$ denotes the Minkowski sum operator: Let $S_1$ and $S_2$ be two sets in $\mathbb{R}^n$: $S_1 \oplus S_2 := \{ x \in \mathbb{R}^n | x = x_1 + x_2, \ x_1 \in S_1, \ x_2 \in S_2 \}$. 

The authors are with Université d’Orléans, PRISME laboratory UPRES EA 4229. 63 av. de Latre de Tassigny, 18000 Bourges, FRANCE. Yasmina.Becis@bourges.univ-orleans.fr

The present work proposes a state estimation approach for models corrupted by bounded disturbances with ellipsoidal noise and state characterizations and complements the one presented in [7] which dealt with multisensor estimation fusion for systems with nonlinear dynamics and with linear-in-state potentially faulty measurements. Here, the multisensor measurements are supposed to be fault-free (cf. Section IV) and a detailed, easy and ready to apply method is presented for the computation of upper bounds, in the sense of positive semi-definiteness, of the Jacobian matrix of the state evolution function which leads to the outer-bounding ellipsoid for the reachable set of this function (cf. Section III). This part was lacking in [7].
II. PROBLEM FORMULATION: TWO-STAGES STATE-BOUNDING PROCESS

Consider a multisensor environment where several sensors observe the same dynamic system, the state vector of which has to be estimated, and where each sensor is attached to a local processor. These processors exchange data with a central processor that computes the state estimate and its enclosing set in light of all the local measurements.

A. Nonlinear state dynamics

The unknown state vector to be estimated, \( x_k \in \mathbb{R}^n \), evolves according to the following discrete-time model

\[
 x_k = \varphi(x_{k-1}, u_{k-1}) + w_{k-1}, \quad k \in \mathbb{N}^*, \tag{1}
\]

where \( u_k \in \mathbb{R}^m \) is a known input vector; \( w_k \in \mathbb{R}^n \) is an unknown bounded additive input (or noise) vector. It may include modeling inaccuracies and discretization errors. The only property it has is

\[
 w_k \in \mathcal{E}(0, W_k) \iff w_k^T W_k^{-1} w_k \leq 1, \tag{2}
\]

where \( W_k \in \mathbb{R}^{n \times n} \) is a known SPD matrix characterizing the shape, the size and the orientation of the ellipsoid containing all possible values of this noise vector; and \( \varphi : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n, (x, u) \to \varphi(x, u) \). For the sake of simplicity, the input \( u_k \) will be omitted by letting \( \varphi_k : \mathbb{R}^n \to \mathbb{R}^n, x \to \varphi_k(x) := \varphi(x, u_k) \). It is assumed that \( \varphi_k \) is a function of class \( C^1 \), that its Jacobian matrix is bounded for any bounded \( x \) and that

\[
 (x_0^* - \hat{x}_0)^T P_0^{-1} (x_0^* - \hat{x}_0) \leq \xi_0^2 \Rightarrow x_0^* \in \mathcal{E}(\hat{x}_0, \xi_0^2 P_0) \tag{3}
\]

where \( \hat{x}_0 \) is the estimate of the initial state vector \( x_0^* \) and where \( P_0 \) (a SPD matrix) and \( \xi_0 \) (an arbitrary positive scalar) are chosen as large as the confidence in \( \hat{x}_0 \) is poor.

At any time step \( k \), the state vector \( x_k \) belongs to the ellipsoid \( \mathcal{E}_k := \mathcal{E}(\hat{x}_k, \xi_k^2 P_k) \) which has to be defined.

B. Multisensor measurements

The system (1) is observed by \( N \) sensors. Each one delivers a measurable system output vector, \( y_{ki} \in \mathbb{R}^{p_i} \),

\[
y_{ki} = F_i x_k + v_{ki}, \quad i \in \{1, \ldots, N\}, \tag{3}
\]

where \( F_i \in \mathbb{R}^{p_i \times n} \), is the output matrix of full row rank and \( v_{ki} \in \mathbb{R}^{p_i} \) is the unobservable measurement noise vector of the \( i^\text{th} \) sensor. It satisfies

\[
v_{ki} \in \mathcal{E}(0, V_{ki}) \iff v_{ki}^T V_k^{-1} v_{ki} \leq 1, \quad \forall k \in \mathbb{N}, \tag{4}
\]

where \( V_i \in \mathbb{R}^{p_i \times p_i} \) is a known SPD matrix characterizing the shape of the ellipsoid containing all acceptable values of this noise vector. The \( i^\text{th} \) measurement equation (3) with the noise bound (4) define an other bounding set for the state vector \( x_k \in \mathcal{S}_{ki} : \)

\[
 \mathcal{S}_{ki} := \{ x \in \mathbb{R}^n | (y_{ki} - F_i x)^T V_k^{-1} (y_{ki} - F_i x) \leq 1 \} \tag{5}
\]

This is a degenerate ellipsoid which projection on the \( \mathbb{R}^{p_i} \) space via the matrix \( F_i^T \) is the ellipsoid \( \mathcal{E}(y_{ki}, V_{ki}) \).

III. TIME UPDATE PROCESS

In this section, we compute the predicted ellipsoid \( \mathcal{E}_{k/k-1} \), which uses the data up to previous step \( k - 1 \), i.e., the ellipsoid \( \mathcal{E}_{k-1} \) containing all possible values of \( \hat{x}_{k-1} \) and the state dynamics (1).

A. Predicted ellipsoid

The following lemma expresses a bounding ellipsoid for \( \varphi_k(x) \), where \( x \in \mathcal{E}(\hat{x}_k, \xi_k^2 P_k) \).

\[
 \text{Lemma 3.1: If } \mathcal{E}_{k-1} = \mathcal{E}(\hat{x}_k, \xi_k^2 P_k) \text{ is bounded then there exists a bounded matrix } \Phi_k^* \in \mathbb{R}^{n \times n}, \text{ such that } \nonumber
\]

\[
 \forall \xi \in \mathcal{E}_{k}, \quad \Phi_k^* \Phi_k^T - \Phi_k(\xi)\Phi_k(\xi)^T \text{ is SPD.} \tag{6}
\]

where \( \Phi_k(\xi) \) is the Jacobian matrix of \( \varphi_k \) at \( \xi : \Phi_k(\xi) := \frac{d\varphi_k(x)}{dx} \), and the following yields

\[
 x \in \mathcal{E}_k \Rightarrow \varphi_k(x) \in \mathcal{E} \left( \varphi_k(\hat{x}_k), \xi_k^2 \Phi_k^* P_k \Phi_k^T \right). \tag{7}
\]

Proof: Let \( \phi : [0, 1] \to \mathbb{R}^n, \tau \to \phi(\tau) = \varphi_k(x_k + \tau(x - \hat{x}_k)) \); then \( \tau \) covers the interval \([0, 1]\), the vector \( x_k + \tau(x - \hat{x}_k) \) sweeps the segment\(^1 \langle x, \hat{x}_k \rangle \). Let us apply the mean value theorem to \( \phi : [0, 1], (\phi(1) - \phi(0)) = \int_0^1 \phi'(\tau) d\tau = \phi(0), \) with \( \phi'(\tau) = \frac{d\varphi_k(x_k + \tau(x - \hat{x}_k))}{dx} \). Since the ellipsoid is convex, \( \forall \xi \in \mathcal{E}_{k} \iff \langle x, \hat{x}_k \rangle \subset \mathcal{E}_{k} \Rightarrow \xi \in \mathcal{E}_{k} \), this leads to

\[
 \exists \xi \in \mathcal{E}_k, \quad \varphi_k(x_k) = \varphi_k(\hat{x}_k) + \Phi_k(\xi)(x - \hat{x}_k). \tag{8}
\]

Since \( \varphi_k \) is of class \( C^1 \) and \( \mathcal{E}_k := \mathcal{E}(\hat{x}_k, \xi_k^2 P_k) \) is bounded, \( \varphi_k \) is locally Lipschitz on \( \mathcal{E}_k \) and there is a matrix \( \Phi_k^* \) meeting the condition (6), which means that \( \forall x \in \mathbb{R}^n, x^T \left[ \Phi_k^* P_k \Phi_k^T \right]^{-1} x \leq \sup_x \tau^T \left( \Phi_k(\xi) P_k \Phi_k(\xi)^T \right)^{-1} \).

Let us, for the sake of simplicity, assume that \( \varphi_k \) is bijective. \( \Phi_k(\xi) \) is then invertible and so is \( \Phi_k^* \). By the use of (8), we can write:

\[
 x \in \mathcal{E}_k \Leftrightarrow (x - \hat{x}_k)^T \Phi_k(\xi)^T P_k \Phi_k(\xi)^T \Phi_k^* \Phi_k^T \leq 1 \nonumber
\]

\[
 \Rightarrow [\varphi_k(x) - \varphi_k(\hat{x}_k)]^T \left[ \Phi_k^* P_k \Phi_k^T \right]^{-1} \left[ \varphi_k(x) - \varphi_k(\hat{x}_k) \right] \nonumber
\]

\[
 \leq [\varphi_k(x) - \varphi_k(\hat{x}_k)]^T \left[ \Phi_k(\xi) P_k \Phi_k(\xi)^T \right]^{-1} \left[ \varphi_k(x) - \varphi_k(\hat{x}_k) \right] \nonumber
\]

\[
 \Rightarrow \varphi_k(x) \in \mathcal{E} \left( \varphi_k(\hat{x}_k), \xi_k^2 \Phi_k^* P_k \Phi_k^T \right). \tag{9}
\]

Theorem 3.2 (Time update / prediction stage): If \( x_{k-1} \in \mathcal{E}(\hat{x}_{k-1}, \xi_{k-1}^2 P_{k-1}) =: \mathcal{E}_{k-1} \) obeying to (1) for any \( w_{k-1} \in \mathcal{E}(0, W_{k-1}) \), then \( \forall \rho \in [0, 1], x_k \in \mathcal{E}(\hat{x}_{k/k-1}, \xi_{k/k-1}^2 P_{k/k-1}) =: \mathcal{E}_{k/k-1} \), where

\[
 \hat{x}_{k/k-1} := \varphi_k(\hat{x}_{k-1}), \tag{9a}
\]

\[
 P_{k/k-1} := \frac{\Phi_k^* - P_{k-1} \Phi_k^T}{1 - \rho \xi_{k-1}^2}, \tag{9b}
\]

\[
 \xi_{k/k-1}^2 := \xi_{k-1}^2. \tag{9c}
\]

\(^1[a, b] (\{a, b\}) := \{ x \in \mathbb{R}^n | x = (1 - \theta)a + \theta b, \theta \in [0, 1] \{\{0, 1]\).\]

\(^2\text{If either } \Phi_k(\xi) \text{ or both } \Phi_k(\xi) \text{ and } \Phi_k^* \text{ are non invertible, their inverses can be replaced, in this proof, by their Moore Penrose pseudo-inverses.}\)
Proof: According to (1), (2) and (7), it is plain to see that $x_k \in \{x \in \mathbb{R}^n | x = \varphi_{k-1}(x_1) + x_2, x_1 \in E_{k-1}, x_2 \in E(0, W_k) \} = E(0, W_k)$, where $\oplus$ denotes the Minkowski sum operator (cf. Section 1, i.e.).

Remark 1 If $\varphi_k$ is not a bijection and/or if a suitable $\Phi_k^*$ is not invertible, the ellipsoid containing all possible values of $\varphi_k(x_k)$, given in (7), is degenerate since its shape matrix $\Phi_k^* P_k \Phi_k^{*T}$ is no longer positive definite. However, as $W_k$ is SPD, $P_{k/k-1}$ is SPD anyway (see (9b)) and the predicted ellipsoid $E_{k/k-1}$ can’t be degenerate.

Remark 2 The parameter $0 < \rho < 1$ can be chosen to minimize, at each time step $k$, either the squared volume of $E_{k/k-1}$, i.e., $s_{k/k-1}^2 = \det P_{k/k-1}$ or the sum of the squared lengths of its axes, i.e., $s_{k/k-1}^2 = \text{tr} P_{k/k-1}$. For the latter case, there is an explicit solution $\rho = \sqrt{\text{tr} W_{k-1}^{-1} (\sqrt{\text{tr} W_{k-1}} + s_{k/k-1}^{-1})}$ (cf. [8]).

B. An upper bound of the Jacobian matrix $\Phi_k$

Here is a proposed method for computing the matrix $\Phi_k^*$ satisfying (6).

Proposition 3.3 ([9]): Given any symmetric matrices $A_c, A_d \in \mathbb{R}^{n \times m}$ (where $A_d$ has non-negative entries), for all symmetric $A \in \mathbb{R}^{n \times m}$, such that $A_c - A_d \preceq A \preceq A_c + A_d$, it holds that

$$\lambda_i(A_c) + \lambda_{\min}(A_d) \leq \lambda_i(A_c + A_d) \leq \lambda_i(A_c) + \lambda_{\max}(A_d),$$

$\forall i \in \{1, \ldots, n\}$.

The proof can be found in [10].

Corollary 3.4: Given any $A_c \in \mathbb{R}^{n \times m}$, $n \geq m$ and any $A_d \in \mathbb{R}^{n \times m}$ with non-negative entries, for all $A \in \mathbb{R}^{n \times m}$, such that $A_c - A_d \preceq A \preceq A_c + A_d$, it holds that

$$|\sigma_i(A) - \sigma_i(A_c)| \leq \sigma_{\max}(A_d), \quad \forall i \in \{1, \ldots, n\}.$$  \hfill (10)

Proof: Applying Proposition 3.3 to the matrices $\begin{pmatrix} 0_n & A^T \end{pmatrix}$, $\begin{pmatrix} 0_n & A^T \end{pmatrix}$ and $\begin{pmatrix} 0_n & A^T \end{pmatrix}$, and recalling that (cf. [11]) $\lambda_i(\begin{pmatrix} 0_n & A^T \\ A & 0_n \end{pmatrix}) \in \{\sigma_1(A), \ldots, \sigma_m(A), 0, \ldots, 0, -\sigma_m(A), \ldots, -\sigma_1(A)\}$, we obtain $|\sigma_i(A_c) - \sigma_i(A_d)| \leq \sigma_i(A_c) \leq \sigma_i(A_c + A_d) + \sigma_{\max}(A_d)$. Other proof: There exists $A' \in \mathbb{R}^{n \times m}$, such that $A = A_c + A'$ and $|A'| \preceq A_d$. This (elementwise) inequality implies that $\sigma_{\max}(A') \leq \sigma_{\max}(A_d)$. From [11], $|\sigma_i(A_c + A') - \sigma_i(A_c)| \leq \sigma_{\max}(A')$ and then $|\sigma_i(A) - \sigma_i(A_c)| \leq \sigma_{\max}(A_d)$.

Theorem 3.5: Let $\Phi_k = \begin{pmatrix} \Phi_{k,i} & \Phi_{k,i} \\ \Phi_{k,i} & \Phi_{k,i} \end{pmatrix}$, $\Psi_k = \begin{pmatrix} \Psi_{k,i} & \Psi_{k,i} \\ \Psi_{k,i} & \Psi_{k,i} \end{pmatrix}$, $i, j \in \{1, \ldots, n\}$ and

$$\Phi_{k,i} := \min_{\xi \in E_k} \phi_{k,i}(\xi), \quad \Phi_{k,i} := \max_{\xi \in E_k} \phi_{k,i}(\xi),$$

$$\Phi_{c_k} := \frac{1}{2} (\Phi_{c_k} + \Phi_{c_k}^T), \quad \Phi_{d_k} := \frac{1}{2} (\Phi_{c_k} - \Phi_{c_k}^T) \quad (11b)$$

and let $\Phi_{c_k} := (\Phi_{c_k}^T \Phi_{c_k})^{1/2}$ be a symmetric matrix. The matrix $\Phi_k^* = \Phi_{c_k} + \|\Phi_{d_k}\|_F I_n$ satisfies the condition (6).

Proof: $\Phi_{c_k} - \Phi_{d_k} \preceq \Phi_k(\xi) \preceq \Phi_{c_k} + \Phi_{d_k}$ and $\Phi_{d_k} \geq 0_{n \times n}$. By virtue of Corollary 3.4, $\forall i = \{1, \ldots, n\}$,

$$\sigma_1(\Phi_k(\xi)) - \sigma_i(\Phi_k(\xi)) \leq \sigma_{\max}(\Phi_{d_k}) \Leftrightarrow \lambda_i(\Phi_k(\xi)) \Phi_k(\xi)) \leq \sigma_{\max}(\Phi_{d_k}) \Leftrightarrow \sigma_i(\Phi_k(\xi)) = \left(\Phi_k(\xi)^2 + 2\sigma_{\max}(\Phi_{d_k}) \sqrt{\lambda_i(\Phi_k(\xi)) \Phi_k(\xi))} \right)^2 = \max(2\sigma_{\max}(\Phi_{d_k}) + \lambda_i(\Phi_k(\xi)) \Phi_k(\xi)) + 2\sigma_{\max}(\Phi_{d_k}) \sqrt{\lambda_i(\Phi_k(\xi)) \Phi_k(\xi))} = \lambda_i(\Phi_k(\xi)) \Phi_k(\xi)) \leq \lambda_i(\Phi_k(\xi)) \Phi_k(\xi)) \leq \lambda_i(\Phi_k(\xi)) \Phi_k(\xi)) \leq \lambda_i(\Phi_k(\xi)) \Phi_k(\xi)) \leq \lambda_i(\Phi_k(\xi)) \Phi_k(\xi)).$$

Observing the prediction mechanism, in particular the equation (9b), it seems to be more efficient to find an upper bound to the symmetric matrix $\Phi_k(\xi) \Phi_k(\xi)^T$ (or even to $\Phi_k(\xi) \Phi_k(\xi)^T$) instead of $\Phi_k(\xi) \Phi_k(\xi)^T$.

Theorem 3.6: Let $\Psi_k(\xi) := \Phi_k(\xi) \Phi_k(\xi)^T$, $\Psi_k = \begin{pmatrix} \Psi_{k,i} & \Psi_{k,i} \\ \Psi_{k,i} & \Psi_{k,i} \end{pmatrix}$, $\Psi_k = \begin{pmatrix} \Psi_{k,i} & \Psi_{k,i} \\ \Psi_{k,i} & \Psi_{k,i} \end{pmatrix}$, $i, j \in \{1, \ldots, n\}$ and

$$\Psi_{k,i} := \min_{\xi \in E_k} \psi_{k,i}(\xi), \quad \Psi_{k,i} := \max_{\xi \in E_k} \psi_{k,i}(\xi),$$

$$\Psi_{c_k} := \frac{1}{2} (\Psi_{c_k} + \Psi_{c_k}^T), \quad \Psi_{d_k} := \frac{1}{2} (\Psi_{c_k} - \Psi_{c_k}^T). \quad (12b)$$

The matrix $\Phi_k^*$ such that $\Phi_k^* \Phi_k^T = \Psi_k := \Psi_{c_k} + \lambda_{\max}(\Psi_{d_k}) I_n$ satisfies the condition (6).

Proof: Proposition 3.3 says that $\lambda_i(\Psi_{k,i}) \leq \lambda_i(\Psi_{c_k}) + \lambda_{\max}(\Psi_{d_k}) I_n = \lambda_i(\Psi_{c_k} + \lambda_{\max}(\Psi_{d_k}) I_n), \forall i = \{1, \ldots, n\}, \Leftrightarrow \Psi_k(\xi), \forall i \in \{1, \ldots, n\} \text{ is SPD}.$

Remarks 3 The use of Theorem 3.6 instead of Theorem 3.5 is done at the expense of some extra calculus : an algebraic expression of a product, $\Phi_k(\xi) \Phi_k(\xi)^T$, of two $n \times n$ matrices, the elements of which are functions of the unknown variable $\xi$ (the optimization argument). Notwithstanding, this results in saving in $n^2 - n$ online optimization processes. Indeed, applying Theorem 3.5, $2n^2$ maximizations/minimizations (11a) have to be performed at each time-step $k$, whereas there are only $n^2 + n$ optimizations (12a) if using Theorem 3.6, thanks to the symmetry of $\Psi_k$ and $\Psi_k$. Besides, when using symmetric matrix bounds, as in Theorem 3.6, one can use the interlacing recursive approach presented in [10]. The reader is also referred to [12] where a filtering method to reduce the overestimation produced by various bounding approaches was developed.

3The Minkowski sum of two ellipsoids can be bounded by a family of parametrized ellipsoids : If $x \in E(x_1, P_1)$ and $x_2 \in E(x_2, P_2)$, then $x_1 + x_2 \in E(x_1, P_1) \oplus E(x_2, P_2) \subseteq E(c_1 + c_2, 1/\rho P_1 + 1 - 1/\rho P_2)$.

4The matrix $\Phi_k^*$ can be, for instance, the Cholesky factorisation of $\Psi_k^*$. 

1885
C. A bound for the Taylor-Lagrange remainder

When the confidence on the state estimate is important enough and the size of the ellipsoid $\mathcal{E}_k$ is not large, bounding the remainder of the first order Taylor series expansion of $\phi_k$ at $\hat{x}_k$ is an alternative solution for bounding the set enclosing all possible values of the images of $x_k$ by the function $\phi_k$. This alternative is however of questionable value since it involves, as shown in the lemma below, tedious calculus with additional $n^3$ (or $n^3 + n^2$), if $\phi_k$ is of class $C_2$ literal function derivatives.

**Lemma 3.7**: If $\phi_k$ is twice differentiable, there exists a matrix $\Delta_k^* \in \mathbb{R}^{n^2 \times n}$ such that $\forall x, \xi_1, \ldots, \xi_n \in \mathcal{E}_k$, 
\[
\phi_k^\ast \Phi_k^T \Delta_k^* \phi_k \in \mathcal{E}_k,
\]
where $\Delta_k^* \in \mathbb{R}^{n^2 \times n}$ is SPDS.

**Theorem 3.5** (as well as Theorem 3.6) can be adapted for the computation of $\Delta_k^*$.

**Theorem 3.8**: Let $\Delta_k = [\Delta_{ij}]$, $\Xi_k = [\Xi_{ij}]$ and $\delta_{k,i} (x, \xi)$ be the $j$th element of the vector $(x - \hat{x}_k)^T \mathcal{H}(\xi_k) (x - \hat{x}_k)$ and $\phi_k (x) - \phi_k (\hat{x}_k)$ be the $k$th component of $\phi_k (x) = (\phi_k (x), \Xi_k) (x - \hat{x}_k)^T + \Delta_k (x, \xi) (x - \hat{x}_k)$. The rest of the proof is similar to the one of Lemma 3.1.

**Remark 4** It is possible to consider the linearization error $\varepsilon_k$ as an additive error (like $w_k$) belonging to an ellipsoid. In fact, for all $x \in \mathcal{E}(\hat{x}_k, \xi_k)$, there exists $\xi$ such that $\varepsilon_k := \psi_k (x) - \phi_k (\hat{x}_k) - \phi_k (\hat{x}_k) (x - \hat{x})$ and $\Delta_k^* = \Delta_k (x, \xi) (x - \hat{x}_k)$.

$\Delta_k^*$ is not necessary SPD. If it is not, the following matrix inverses are replaced by the Moore-Penrose pseudo-inverses. This matrix can be computed exactly at the same way as $\Phi^*$ in Theorems 3.5 and 3.6.
Proposition 4.2 ([13]): If $F_{ki}$ has full row rank then the optimization problem 
$$
\min_{\omega_{ki} \in \mathbb{R}^+} \max_{v_{ki} \in E(0, V_{ki})} V_{ki} \left( x_k - \hat{x}_{k-1} \right),
$$
which is equivalent to 
$$
\min_{\omega_{ki} \in \mathbb{R}^+} \max_{v_{ki} \in E(0, V_{ki})} V_{ki}, \quad \text{if } \omega_{ki} \in \mathbb{R}^+ \text{ and } v_{ki} \in E(0, V_{ki}),
$$
then replacing $P_{k/k-1}^{1}$ by $P_{k/k-1}$ and $\hat{x}_{k/k-1}^{1}$ by $\hat{x}_{k/k-1}$, then (18a) follows.

Remark 5
The value of $\omega$ given by Proposition 4.2 leads to 
$$
\left\| (I - \omega_{ki}, F_{ki}, F_{ki}^T V_{ki}) \right\|^2_{V_{ki}^{-1}} = 1 \quad \text{which is equivalent to } 
\left\| \omega_{ki} - F_{ki}, \hat{x}_{ki} \right\|^2_{V_{ki}^{-1}} = 1 \quad \text{which means that the a posteriori output error vector belongs to the output noise ellipsoid } \mathcal{E}(0, V_{ki}).
$$

Remark 6
The local processors are performing their local estimates in an open-loop manner. Indeed, the $N$ local state estimates (both predictions and corrections) are independently computed at the local processors.

Firstly, this (pseudo) solution is somewhat computationally heavy. Secondly, it produces $N$ possibly different state estimates and thirdly, the processors don’t communicate with each other, thus this approach doesn’t take any advantage from the multi-sensor framework.

Theorem 4.1 and Proposition 4.2 are not useless for all that, they will be used in the sequel.

B. The global estimation at the fusion agent

Since the prediction stage is performed without the need for measurements, it should be done at the central processor once instead of $N$ times when tackled by local agents (as in previous subsection), but then it needs the global state estimate $\hat{x}_{k-1}$ instead of the local one $\hat{x}_{k-1}$:

Theorem 4.3 (Global observation update): Let $x_k \in \mathcal{E}_{k/k-1}$, let $y_{ki}$ be given by (3) for any $v_{ki}$ satisfying (4) and let

$$
\hat{x}_k := \hat{x}_{k/k-1} + P_k \sum_{i=1}^{N} \omega_{ki} F_{ki}^T V_{ki}^{-1} \delta_{ki},
$$
(17a)

$$
P_k := P_{k/k-1}^{-1} + \sum_{i=1}^{N} \omega_{ki} F_{ki}^T V_{ki}^{-1} F_{ki},
$$
(17b)

$$
\omega_{ki} := \frac{1}{\omega_{ki} \in \mathbb{R}^+}, \quad \forall i \in \{1, \ldots, N\},
$$
(17c)

$$
\delta_{ki} := y_{ki} - F_{ki}, \hat{x}_{k/k-1},
$$
(17d)

then

$$
\forall \omega_{ki}, \ldots, \omega_{kN} \in \mathbb{R}^+, \quad x_k \in \mathcal{E}_{k/k-1} \cap S_{k_1} \cap \ldots \cap S_{k_N} \subseteq \mathcal{E}(\hat{x}_k, \omega_{k1}^2 P_{k/k-1}) =: \mathcal{E}_k,
$$
then $S_{k_1}$ and $\mathcal{E}_{k/k-1}$, are given by (5) and Theorem 3.2 respectively.

Proof: Considering that if $x \in \bigcap_{i=1}^{N} S_{k_i} \cap \mathcal{E}_{k/k-1}$ then 
$$
(x - \hat{x}_{k/k-1})^T P_{k/k-1}^{-1} (x - \hat{x}_{k/k-1}) + \sum_{i=1}^{N} \omega_{ki} (y_{ki} - F_{ki}, x)^T V_{ki}^{-1} (y_{ki} - F_{ki}, x) \leq \omega_{ki}^2 \leq \omega_{ki}^2, \quad \forall \omega_{ki} \in \mathbb{R}^+.
$$

In Theorem 4.3, the global state estimate and its bounding ellipsoid are computed using all local data collected by the $N$ sensors. These measurements and noise characteristics have to be transmitted to the central processor.

However, it is possible for this processor to compute the global ellipsoid $\mathcal{E}(\hat{x}_k, \omega_{k1}^2 P_{k/k-1})$ by using only the local ellipsoids $\mathcal{E}(\hat{x}_{ki}, \omega_{ki}^2 P_{ki})$ (in addition to the predicted one, obviously), obtained at the local processors after the correction stage, according to Theorem 4.1 and Proposition 4.2, and transmitted to the central processor. The latter has no need to receive other information than the attributes of these local ellipsoids $\mathcal{E}(\hat{x}_{ki}, P_{ki} \omega_{ki}^2, \omega_{ki} \in \mathbb{R}^+)$ from local agents (neither measurements $y_{ki}$ nor noises characterizations $V_{ki}$).

In this purpose, the quantities $\hat{x}_k, P_k$ and $\omega_{k1}^2$ defining the global bounding ellipsoid for the state vector are expressed as functions of $\hat{x}_{ki}, P_{ki}$ and $\omega_{ki}^2$; $\hat{\omega}_{ki/k-1}$, $\hat{x}_{k/k-1}, P_{k/k-1}$ and $\omega_{k1/k-1}$ are defined in (15), where $\hat{x}_{k/k-1}, P_{k/k-1}$ and $\omega_{k1/k-1}$ are replaced by $\hat{x}_{ki/k-1}, P_{ki/k-1}$ and $\omega_{ki/k-1}$, respectively, which are given in (9). Then

$$
\hat{x}_k := P_k \left( \sum_{i=1}^{N} P_{ki/k-1} \hat{x}_{ki} - (N-1)Q_{k-1/k-1} \hat{x}_{k/k-1} \right),
$$
(18a)

$$
P_k := P_{k/k-1} - (N-1) \sum_{i=1}^{N} P_{ki/k-1},
$$
(18b)

$$
\omega_{k1/k-1}^2 := \left( \sum_{i=1}^{N} \omega_{ki}^2 \right) - \sum_{i=1}^{N} \omega_{ki} (y_{ki} - F_{ki}, x)^T V_{ki}^{-1} (y_{ki} - F_{ki}, x),
$$
(18c)

Proof: From (17a) and (15a), it comes that 
$$
P_{k/k-1}(\hat{x}_k - \hat{x}_{k/k-1}) = \sum_{i=1}^{N} \omega_{ki} F_{ki}^T V_{ki}^{-1} \delta_{ki},
$$
then 
$$
P_{k/k-1}(\hat{x}_k - \hat{x}_{k/k-1}) = \sum_{i=1}^{N} \omega_{ki} F_{ki}^T V_{ki}^{-1} \delta_{ki},
$$
by $P_{k/k-1}$ and $\hat{x}_{k/k-1}$, by $\hat{x}_{k/k-1}$, (18a) follows.
In the same way, from (17b) and (15b), we have
\[ P_{k-1}^{-1} - P_{k-1/k-1}^{-1} = \sum_{i=1}^{N} \omega_{ki} F_{ki} V_{ki}^{-1} F_{ki}^{T} \]
Then replacing $P_{k-1/k-1}$ by $P_{k-1/k-1}$, (18b) follows. Now, by virtue of (17c) and (15c), we can write
\[ \sum_{i=1}^{N} \omega_{ki} (1 - \delta_{ki}) T V_{ki}^{-1} \delta_{ki} = - \sum_{i=1}^{N} (\hat{x}_{ki} - \hat{x}_{k/k-1})^{T} P_{k-1}^{-1} (\hat{x}_{ki} - \hat{x}_{k/k-1}) + \sum_{i=1}^{N} \left( \gamma_{k-1}^{2} - \gamma_{k-1/k-1}^{2} \right). \]
Replacing $\hat{x}_{k/k-1}$ by $\hat{x}_{k/k-1}$ and $\gamma_{k/k-1}^{2}$ by $\gamma_{k/k-1}$, (18c) follows.

The following important proposition shows that the values of the parameters $\omega_{ki}$ that minimizes the global estimation error norm in the worst noise case are still computed separately (and in parallel) by the local agents.

**Proposition 4.5:** If all $F_{ki}$ have full row rank then the solution to the optimization problem

\[ \min_{\omega_{k1}, \ldots, \omega_{kN} \in \mathbb{R}^{+}} \max_{v_{k1}, \ldots, v_{kN} \in \mathcal{E}(0, V_{ki})} \gamma_{k}, \quad \gamma_{k} := (x_{k} - \hat{x}_{k})^{T} P_{k}^{-1} (x_{k} - \hat{x}_{k}), \]

is obtained by deriving $\gamma_{k}^{-1}$ given by (16) for $i \in \{1, \ldots, N\}$ where $\delta_{ki}$ is defined in (17d).

**Proof:** This result is obtained by deriving $\gamma_{k}^{-1}$ given by (17), after some reformulations, with respect to $\omega_{k1}, \ldots, \omega_{kN} \in \mathbb{R}^{+}$.

Let us recapitulate. At each time step $k$, the fused data of the preceding time step, $\hat{x}_{k-1}$, $P_{k-1}$ and $\gamma_{k-1}^{2}$ are used to compute the predicted ellipsoid by the central processor, according to *Theorem 3.2* and other theorems of Section III. Then $\hat{x}_{k/k-1}$, $P_{k/k-1}$ and $\gamma_{k/k-1}^{2}$ are returned to the local agents where they are used instead of $\hat{x}_{k/k-1}$, $P_{k/k-1}$, and $\gamma_{k/k-1}^{2}$, to perform the observation updating producing $\hat{x}_{k}$, $P_{k}$ and $\gamma_{k}^{2}$, by the means of *Theorem 4.1* and *Proposition 4.2*. Finally, *Theorem 4.4* is used to deduce from these local data the fused ones for the time step $k$: $\hat{x}_{k}, P_{k}$ and $\gamma_{k}^{2}$.

**V. CONCLUSION**

A hierarchical set-membership estimation method is presented where different sensors offer several measurements of the same state vector, each of them subject to sensor noise vector belonging to known ellipsoid.

The prediction (or time update) stage is performed by computing the ellipsoidal superset for the Minkowski sum of the unknown input ellipsoid and the one containing the nonlinear transformation of the state ellipsoid by the function $\varphi_{k}$.

The correction (or observation update) stage is performed on two levels. At the lower level, each local state estimate and its outer-bounding ellipsoid is computed separately using the measurements given by the corresponding sensor. All the local results are transmitted to the central processor which computes the global estimate with its bounding ellipsoid and feeds this information back to all the local processors which update their data using this global estimate.

**The overall algorithm for the global state estimation:**

1. $k \leftarrow 1$.
2. The predicted ellipsoid $\mathcal{E}(\hat{x}_{k/k-1}, \gamma_{k/k-1}^{2} P_{k/k-1})$ (which depends only on the system dynamics $(\varphi_{k}, \Phi_{k-1}^{*})$, the process noise characterizations $W_{k-1}$ and $\mathcal{E}(\hat{x}_{k-1}, \gamma_{k-1}^{2} P_{k-1})$) is computed by the central processor according to (9) with an appropriate choice of the parameter $\rho_{k}$ (see the Remark III-A) and where $\Phi_{k-1}^{*}$ is computed according to *Theorem 3.5*, 3.6 or 3.8. Then $\hat{x}_{k/k-1}$ is transmitted to the local processors setting $\hat{x}_{k/k-1} \leftarrow \hat{x}_{k/k-1}, P_{k/k-1} \leftarrow P_{k/k-1}$ and $\gamma_{k/k-1}^{2} \leftarrow \gamma_{k/k-1}^{2}$.
3. The local processors compute the corrected ellipsoids $\mathcal{E}(\hat{x}_{k}, \gamma_{k}^{2} P_{k})$ using their respective measurements, $y_{ki}$, and $\gamma_{k}^{2}$, the output noise characterizations, $V_{ki}$, and $\mathcal{E}_{k/k-1}$ according to (15) and *Proposition 4.2*.
4. The local processors transmit these ellipsoids $(\mathcal{E}_{k})$ to the fusion center which deduces, using $\hat{x}_{k/k-1}$, the unique global ellipsoid $\mathcal{E}(\hat{x}_{k}, \gamma_{k}^{2} P_{k})$ containing all possible values of the state vector $x_{k}$ according to (18).
5. $k \leftarrow k + 1$ then go to 2.

**REFERENCES**


