AN APPROACH FOR DISTRIBUTED KALMAN FILTERING

Rogério Bastos Quirino & Celso Pascoli Bottura
DMCSI/FEEC/UNICAMP
13083-970 – Campinas – SP – Brasil
rquirino@dmcsi.fee.unicamp.br certto.com.br
cpb@turing.unicamp.br

Abstract: In this article we propose a parallel and distributed state estimation structure developed from an hierarchical estimation structure, optimal in the sense of Kalman filtering and that is based on the multiple projections method. We explore a duality that exists between two state space representations, derived from the application of an approach based on the coupling and noise terms of the original system. The algebraic structure developed is suboptimal, due to the fact that it does not take into account the corrections of the state predictions based on the multiple innovations. This approach contributes to the design of distributed Kalman filtering algorithms.

1 INTRODUCTION

In distributed estimation problems, parallelism arises naturally due to data obtained by different subsystems located at geographically dispersed locations.

Among the solutions based on hierarchical techniques are those proposed in (Hassan *et alii*, 1978; Hashemipour, Roy, and Laub, 1987), where the existence of information exchange between levels and between the subsystems at the same level depends on the mathematical development utilized to decompose the Kalman filter.

A computational procedure to transform an hierarchical Kalman filter into a partially decentralized estimation structure is developed in (Quirino, Bottura & Costa Filho, 1998).

In order to contribute to the study and analysis of construction strategies for distributed estimation algorithms, we propose a distributed filter developed from an hierarchical filter applied to an approximate state space representation. This approximate state space representation is based on the replacement of the interconnection and noise terms of the state and output equations by white noise processes. The concept of innovation processes is used to show that these interconnection and noise terms can be approximated as white noise processes, particularly in the important practical case of low measurement noise systems.

Moreover the derivation reveals that a good approximation results for the weakly coupled systems case.

Artigo Submetido em 04/01/99
1a. Revisão em 28/06/99; 2a. Revisão em 17/04/00; 3a. Revisão em 07/11/00.
Aceito sob recomendação do Ed. Consultor Prof. Dr. Liu Hsu

In section 2 we present the hierarchical filter obtained in (Hassan *et alii*, 1978). applied to the approximate system.

In section 3 we develop the new filter based on a theorem that provides the sufficiency conditions to yield it.

In addition, an execution diagram which illustrates its implementation is presented.

In section 4 the computational advantage of the new filter is examined.

In section 5 the new filter is simulated for a discrete model of a section of the river Cam near Cambridge in England.

Finally in section 6 the results are compared with those for the Hierarchical Kalman Filter (**HKF**) and Totally Decoupled Hierarchical Kalman Filter (**TDHKF**) solutions.

2 FORMULATION OF THE PROBLEM

Consider the linear discrete-time system at instant k+1, comprising N interconnected dynamic linear subsystems, defined by:

$$x_{k+1}^i = A_k^i x_k^i + \sum_{\substack{j=1 \\ i \neq j}}^{S} A_k^{ij} x_k^j + w_k^i$$
$$y_{k+1}^i = H_{k+1}^i x_{k+1}^i + v_{k+1}^i + H_{k+1}^{ij} x_{k+1}^j \tag{1}$$

where

x_k^i	i-th subsystem state vector
A_k^i, A_k^{ij}	state transition matrices
w_k^i, v_k^i	zero mean independent Gaussian white noise vectors
y_k	measurement vector
H_k^i, H_k^{ij}	observation matrices

i ,j =1,2,…,s , s≤N

The noise covariances are known and given by

$$E(w_j.w_k^t)=Q_k.\delta j_k,\ E(v_j.v_k^t)=R_{jk}.\delta_{jk}$$

where

δ_{jk} is the kronecker delta function

E is the expected value

Q and R are block-diagonal matrices

Consider the following approximations (Shah, 1971) on the noise terms of system (1):

$$w^{*i}_k = \sum_{\substack{j=1\\ i\neq j}}^{S} A^{ij}_k \tilde{x}^j_{k/k} + w^i_k$$
$$v^{*i}_k = v^i_k + \sum_{\substack{j=1\\ i\neq j}}^{S} H^{ij}_k \tilde{x}^j_{k/k-1} \tag{2}$$

where

$\tilde{x}^j_{k/k}, \tilde{x}^j_{k/k-1}$ are the j-th subsystem prediction and estimation error vectors

Partial discussions on the approximate system disturbance w^{*i}_k and on the approximate measurement disturbance v^{*i}_k as white noise sequences of known covariances are given in Appendix A.

In the following the main results of our application using that approach are described.

Applying the approximations w^{*i}_k and v^{*i}_k to system (1), the following approximate state space representation results:

$$x^i_{k+1} = A^i_k x^i_k + \sum_{\substack{j=1\\ i\neq j}}^{S} A^{ij}_k \hat{x}^j_{k/k} + w^{*i}_k$$
$$y^i_{k+1} = H^i_k x^i_{k+1} + v^{*i}_{k+1} + \sum_{\substack{j=1\\ i\neq j}}^{S} H^{ij}_{k+1} \hat{x}^j_{k+1/k} \tag{3}$$

Let us consider the state estimation using the approximate system proposed in (3), and the hierarchical structure proposed in (Hassan *et alii*, 1978). , including its notation. Part of the resulting equations of this development, the prediction stage of the Kalman filter, is given by equations (4) and (5):

<u>Prediction Equations</u>

$$\hat{x}^i_{k+1/k} = A^i_k \hat{x}^i_{k/k} + \sum_{\substack{i=j\\ i\neq j}}^{N} A^{ij}_k \hat{x}^j_{k/k} \tag{4}$$

$$P^i_{k+1/k} = \alpha^i_{k+1/k}.A^{i\,t}_k + Q^{*i}_k \tag{5}$$

where

$$\alpha^i_{k+1/k} = A^i_k P^i_{k/k} \tag{6}$$

$$Q^{*i}_k = Q^i_k + \alpha^{ij}_{k+1/k} \tag{7}$$

$$\alpha^{ij}_{k+1/k} = \sum_{\substack{i=1\\ i\neq j}}^{S} A^{ij}_k P^j_{k/k} A^{ij\,t}_k \tag{8}$$

Q^{*i}_k is the w^{*i}_k white noise covariance matrix;

$P^i_{k/k}$ is the i-th subsystem covariance matrix .

It is important to observe that the application of this supplemented white noise approach to the original system (1) allows decoupled calculations in the estimation process via the diagonalization of the error covariance matrices $P_{k/k}$ and $P_{k+1/k}$.

The sufficiency conditions to yield the correction equations are established through Theorem I in the next section.

3 PARTIALLY DECOUPLED HIERARCHICAL KALMAN FILTER (PDHKF)

The diagonalization of the approximate state error covariance matrices leads us to the exploitation of the naturally existing parallelism to calculate equations (3)-(5).

For a system partitioned in two subsystems i,j, the parallel calculations of the equations (3)-(5) are depicted in scenarios 1 and 2 of Fig.1.

Note that the algebraic structure resulting from the application of this approach allows us to execute the calculations in a partially decentralized manner, contrary to the multilevel algebraic structure proposed in (Hassan *et alii*, 1978). , which requires the coordination module in order to calculate the off-diagonal terms in the estimation process.

At the recursion level $\ell=1$ of the algorithm proposed in (Hassan *et alii*, 1978). , we can write the correction equation based on the local measurements, in the following form:

$$\hat{x}^i_{(k+1/k+1)_1} = \hat{x}^i_{k+1/k} + G^0_{i1(k+1)}.\tilde{y}^i_{k+1/k} \tag{9}$$

where

$$G^0_{i1(k+1)} = P^i_{k+1/k}.H^{i\,t}_1.P^{y_i\,-1}_{k+1/k} \tag{10}$$

$$\tilde{y}^i_{k+1/k} = H^i_{k+1}\tilde{x}^i_{k+1/k} + v^{*i}_{k+1} \tag{11}$$

From (11), we can write :

$$P^{y_i}_{k+1/k} = H^i_{k+1}.P^i_{k+1/k}.H^{i\,t}_{k+1} + R^{*i}_{k+1} \tag{12}$$

Replacing (11) into (9), results in:

$$\hat{x}^i_{(k+1/k+1)_1} = \hat{x}^i_{k+1/k} + G^0_{i1(k+1)}.(y^i_{k+1} - H^i_{k+1}\hat{x}^i_{k+1/k} - \sum_{\substack{j=1\\ i\neq j}}^{S} H^{ij}_{k+1}\hat{x}^j_{k+1/k}) \tag{13}$$

$$G^0_{i1(k+1)} = P^i_{k+1/k}.H^{i\,t}_1(.H^i_{k+1}.P^i_{k+1/k}.H^{i\,t}_{k+1} + R^{*i}_{k+1})^{-1} \tag{14}$$

where

The covariance matrix of $\tilde{x}^i_{(k+1/k+1)_1}$, based on y^i_{k+1}, can be written as:

$$P^i_{(k+1/k+1)_1} = G^i_{k+1}.P^i_{k+1/k} \tag{15}$$

where

$$G^i_{k+1} = I - G^0_{i1(k+1)}.H^i_{k+1} \tag{16}$$

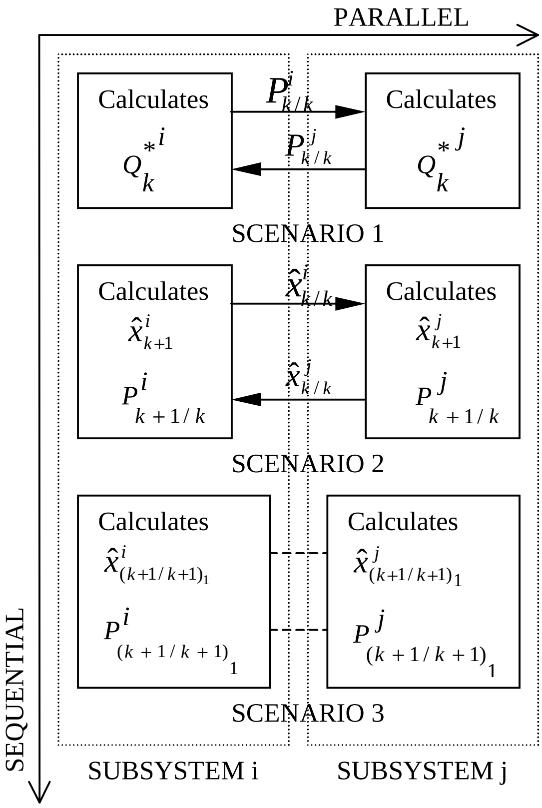

Fig. 1 – Execution Scenarios for the Distributed Filter.

The operation execution for equations (13) - (16) can be visualized in scenario 3 of Fig.1, where no communication among the local subsystems is necessary, as opposed to the original hierarchical structure proposed in (Hassan *et alii*, 1978). .

In scenario 3 of Fig. 1, we complete the distributed calculation of the prediction correction stage based on the local observations.

The off-diagonal covariance matrix for the state estimation error at level $\ell = 1$ is given by:

$$P^{ij}_{(k+1/k+1)_1} = G^{i}_{k+1}.P^{ij}_{k+1/k}.G^{j\,t}_{k+1} \qquad (17)$$

The off-diagonal covariance matrix for the output prediction error is:

$$P^{y_{ij}}_{(k+1/k)} = H^{i}_{k+1}.P^{ij}_{k+1/k}.H^{j\,t}_{k+1} \qquad (18)$$

The non local innovations $\tilde{y}^{i}_{j\,(k+1/k+1)}$, described by:

$$\tilde{y}^{i}_{j\,(k+1/k+1)} = y^{j}_{k+1} - E\,(y^{j}_{k+1} / \mathrm{Y}_k , y^{i}_{k+1}) \qquad (19),$$

and derived by the multiple orthogonalizations

$\boldsymbol{E}\,(\boldsymbol{x}^{i}_{k+1} / \mathrm{Y}_{\boldsymbol{k}} , \boldsymbol{y}^{1}_{k+1} , \boldsymbol{y}^{2}_{k+1} ,\ldots, \boldsymbol{y}^{N}_{k+1})$,allow us to write the filtered state correction as:

$$\hat{x}^{i}_{(k+1/k+1)} = \hat{x}^{i}_{(k+1/k+1)_1} + G^{ij}_{k+1}.\tilde{y}^{i}_{j(k+1/k+1)} \qquad (20)$$

where

$$G^{ij}_{k+1} = G^{i}_{k+1}.P^{ij}_{k+1/k}.H^{j\,t}_{k+1}.P^{\tilde{y}^{i}_{j}\,-1}_{k+1/k+1} \qquad (21)$$

Theorem I:

The state estimation of the aproximate system (3), based on the multiple projections method (Hassan *et alii*, 1978). , generates a partially decoupled distributed filter structure, described by the following distributed estimation equations:

$$\hat{x}^{i}_{(k+1/k+1)} = \hat{x}^{i}_{(k+1/k+1)_1} \qquad (22)$$

$$P^{i}_{(k+1/k+1)} = P^{i}_{(k+1/k+1)_1} \qquad (23)$$

Proof:

From the duality that exists between the state space representations (1), (3) and from the calculations of $P^{i}_{k/k}$ and $P^{ij}_{k+1/k}$, for both representations, the following equations must be satisfied:

$$A^{ij}_{k} P^{ij\,t}_{k/k} A^{ii\,t}_{k} = -A^{ii}_{k} P^{ij}_{k/k} A^{ij\,t}_{k} \qquad (24)$$

$$P^{ij}_{k+1/k} = A^{ii}_{k} P^{ij}_{k/k} A^{jj\,t}_{k} \qquad (25)$$

$$\forall A_k$$

To satisfy equation (24) a sufficient condition is that the cross-covariance sub-matrices of $P_{k/k}$ for the global system be zero, i.e.,

$$P^{ij}_{k/k} = P^{ij\,t}_{k/k} = 0 \qquad (26)$$

Replacing (26) into (25), results in :

$$P^{ij}_{k+1/k} = 0 \qquad (27)$$

Applying equation (27) to the equations (17) and (18), results in:

$$P^{y_{ij}}_{(k+1/k)} = P^{ij}_{(k+1/k+1)_1} = 0 \qquad (28)$$

This is one of the reasons why the coordination module from (Hassan *et alii*, 1978). is not necessary in our proposal.

Then, as a consequence, $P_{k/k}$ and $P_{k+1/k}$ for the approximate system will be diagonal and the gain described by equation (21) will be zero. Q.E.D.

The hypothesis (26) would be a reasonable assumption in cases where the correlations between the subsystems are weak.

If we assume the sufficient condition (26) , we can treat the local observations of the overall system as strictly local observations corrupted by white noise sequences of known covariances. The utilization of this assumption leads us to not consider the correction of the state prediction based on the multiple orthogonalizations on the partitioned measurement vector of the overall system. This characterizes the suboptimality of our proposal.

From a practical point of view, this fact would reduce drastically the communication and synchronization requirements that exist in the hierarchical structure filter(Hassan *et alii*, 1978). when used in distributed processing environments (Quirino, Bottura & Costa Filho, 1998).

Therefore, from equations (22) and (23), we can state that the estimation process is completed in a distributed form as represented in scenario 3 of Fig.1.

4 COMPUTATIONAL REQUIREMENTS

The computational requirements of the new filter as compared to those of the hierarchical Kalman filter can be divided into four categories, i.e. , storage requirements, communication and synchronization time requirements, computational time requirements and quantity of shared data requirements. The storage requirements and the number of shared data of the new filter are significantly smaller than those of the HKF.

The computational time requirements and the communication and synchronization time requirements of the new filter require further elaboration.

Furthermore, good measures of computational time requirements for the HKF and for the new filter are given by the number of elementary multiplication operations involved.

We first consider the number of elementary multiplication operations required for the HKF.

4a) Number of multiplications required for the HKF

Assuming that the partitioned sets of measurements are completely uncoupled, that each subsystem has the same number of state variable n/N and an equal number of measurements m/N, where N is the number of subsystems, x is of dimension n, y is of dimension m, then the number of multiplications required by the HKF is:

$$1.5n^2 + 1.5n^3 + N\left\{\frac{nm}{N^2} + \frac{nm(m+N)}{2N^3} + \frac{m^2\left(\frac{3m}{N}+1\right)}{2N^2}\right\}$$
$$+ N^2\left\{\frac{n^2}{N^3} + \frac{nm^2}{N^3} + \frac{nm}{N^2} + \frac{nm(n+N))}{2N^3}\right\}$$
$$+ \frac{N^2(N-1)}{2}.\frac{n^2 m}{N^3}$$

4b) Number of multiplications required for the new filter

Assuming all the conditions of item 4a), the number of multiplications required by the new filter is:

$$N\left[\frac{6n^3}{N^3} + \frac{2n^2}{N^2} + \frac{nm}{N^2}\left(\frac{3n}{N} + \frac{2m}{N} + 2\right)\right]$$

In a similar way as shown in (Quirino, Bottura & Costa Filho, 1998), we can show that for high order systems, the new filter will give substantial savings in computational time.

5 EXAMPLE

In this section, a practical example is used to illustrate the new approach. In this example the objective is to estimate the states of a two reaches discrete-time model of the river Cam near Cambridge, developed in (Tamura, 1974) .

The system is assumed to be drawn from a zero mean Gaussian white noise process; it is also assumed that the second and fourth states of the system, which represent the amount of dissolved oxygen(DO) in the stream of the first and second reaches, are measured. The model is given by:

$$\begin{pmatrix} z^1_{k+1} \\ q^1_{k+1} \\ z^2_{k+1} \\ q^2_{k+1} \end{pmatrix} = \begin{pmatrix} 0.18 & 0 & 0 & 0 \\ -0.25 & 0.27 & 0 & 0 \\ 0.55 & 0 & 0.18 & 0 \\ 0 & 0.55 & -0.25 & 0.27 \end{pmatrix} \begin{pmatrix} z^1_k \\ q^1_k \\ z^2_k \\ q^2_k \end{pmatrix} + \begin{pmatrix} w^1_k \\ w^2_k \\ w^3_k \\ w^4_k \end{pmatrix} + \begin{pmatrix} 4.5 \\ 6.15 \\ 2.0 \\ 2.65 \end{pmatrix}$$

$$\begin{pmatrix} y^1_{K+1} \\ y^2_{K+1} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} z^1_{k+1} \\ q^1_{k+1} \\ z^2_{k+1} \\ q^2_{k+1} \end{pmatrix} + \begin{pmatrix} v^1_{K+1} \\ v^2_{K+1} \end{pmatrix}$$

where z_1 and z_2 represent the biochemical oxygen demand (BOD) for the first and second reaches, respectively.

In a typical test, the overall system noise Q was $5I_4$, where I_4 is a fourth order identity matrix. The overall measurement noise was $R=I_2$, and the overall covariance matrix of the initial state X_0 was $P_0=25I_4$.

For comparison, the above system was solved using the hierarchical Kalman filter equations(Hassan *et alii*, 1978). , the hierarchical Kalman filter with zero coupling (totally decoupled, i.e., the system coupling terms between the two second order systems are ignored), and the approach developed here with the measurement noise levels of orders of 20%, 30% , 40% and 50% .

The same noise sample functions were used for the three filters, and they were simulated over a period of 100 sample time points.

The noise to signal ratio(NSR) was computed and the results are shown in Table 1.

6 FIGURES OF MERIT AND RESULTS

In this section we propose two figures of merit and discuss some results.

6a) Figures of Merit and Estimates

For the evaluation of the state estimation for each state X_j , for j=1,...,4 , we define the following State Figure of Merit SFM :

State figure of merit

$$SFM(X^j) = \sum_{k=1}^{N} E\left(X^j_k - \hat{X}^j_{k/k}\right)^2$$

For the evaluation of the global quality of the state estimations we define the Total Figure of Merit TFM:

$$TFM = \sum_{j=1}^{4}\sum_{k=1}^{N} E\left(X^j_k - \hat{X}^j_{k/k}\right)^2$$

Results on SFM and TFM, together with the relative noise levels and measurements noise covariance matrices, are presented in Table 1.

The results in Table 1 show that the performance of PDHKF is very close to the HKF, whereas the performance of the zero coupling HKF deteriorates when the noise to signal ratio increases.

Since the measurements are uncoupled, sequential processing is not used to process y^1 , y^2 for both of the partitioned subsystems.

The real value of the state X_3 that is not directly observable through the output and its estimates with the three filters are shown in Fig.2.

The HKF filter shows that a good estimation around the mean value of the real state is obtained. This arises from the fact that initial values of the state and error covariance matrices were known.

Also, in Fig.2, the TDHKF response shows that the state estimate has an error, i.e., it does not preserve the unbias property. In this case, for practical purposes the estimate is not good enough. The error results from ignoring the system's coupling term A_{21} .

In a practical system there will be a large number of subsystems and hence the overall effect of ignoring the coupling terms may be much more serious than that encountered in this two reaches discrete model.

Moreover, in Fig.2 , the PDHKF response shows that the estimation of X_3 is identical to the HKF concerning the expected value.

For X_4 it appears that the results of the TDHKF and PDHKF are better than the HKF. However, it is believed that this may be due to the presence of numerical errors in the computation.

Also the occurrence of numerical errors in the estimation of the states X_1 and X_2 , is attributed to the fact that we obtained two different results for the three filter implementations, when theoretically these results should be exactly identical, due to the inherent decoupling in the model with $A_{12}=0$.

6b) Computational Considerations

For the case of this two reaches discrete model the computational requirements are:

HKF: 464 multiplications per iteration
PDHKF: 152 multiplications per iteration
TDHKF: 60 multiplications per iteration

Hence the PDHKF filter leads to a substantial saving in the number of computations required at each iteration and this in turn leads to a proportional saving in computational time, with good state estimation results.

7 CONCLUDING REMARKS

The central feature of the partial decoupling approach presented here is that it uses the partitioning philosophy to overcome the computational difficulties encountered in the implementation of hierarchical Kalman filtering theory and at the same time accounts for the interaction between the subsystems.

In most industrial systems, the overall plant is composed of interconnected subsystems and this partial decoupling approach gives the designer the freedom to deal with interconnected systems separately.

In the derivation of the approximate system measurement noise covariances, the criterion for partitioning must be based on weak interactions between the subsystems. It remains to show that this partitioning method is applicable to systems where interactions may be strong.

Therefore, application studies for this partitioning approach to more complex large-scale systems must assure that its use leads to a large computational saving without loss of accuracy of the state estimates.

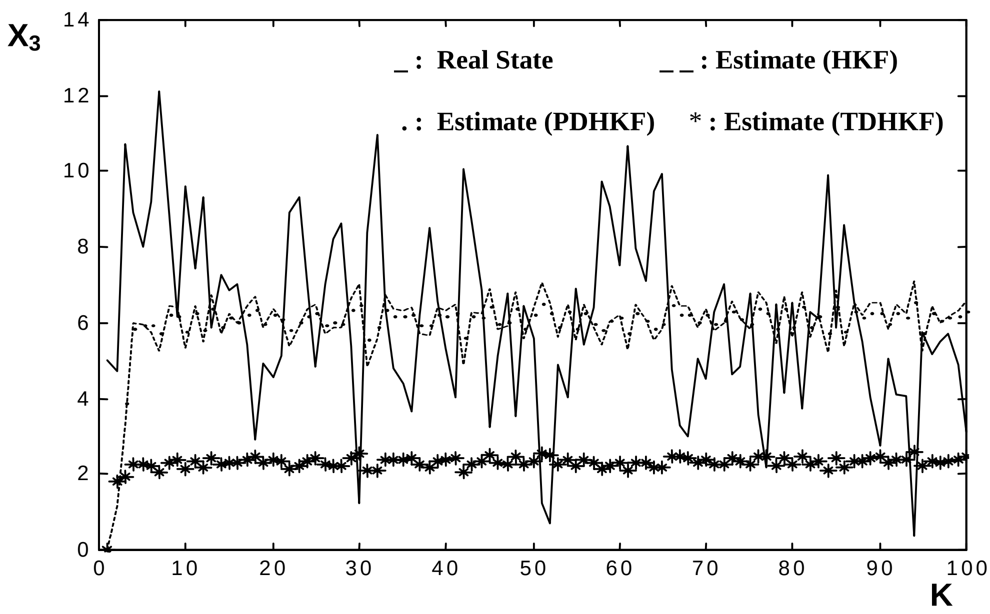

Fig. 2 Estimation of the State X_3 using HKF **,** PDHKF **and** TDHKF
(63% measurement noise level)

Table 1 – Table of Figure of Merit for various noise to signal levels using HKF , PDHKF **and** TDHKF **Filters.**

Filter	Percentage Noise to Signal Ratio		Measurement Noise Covariance Matrix R	Figures of Merit				
				SFM (X^j)				TFM
	Y^1	Y^2		X_1	X_2	X_3	X_4	
HKF	29%	46%	$(5x10^{-4})$ x I_2	4.1724	0.2503	7.4354	0.2503	12.1084
PDHKF				4.1783	0.2503	10.5480	0.2500	15.2266
TDHKF				4.1783	0.2503	19.6260	0.2500	24.3046
HKF	36%	53%	$(5x10^{-2})$ x I_2	5.8774	0.2906	10.231	0.2910	16.6900
PDHKF				5.8845	0.2906	14.6560	0.2860	27.2922
TDHKF				5.8845	0.2906	29.6710	0.2890	36.1351
HKF	46%	63%	I_2	6.6534	0.5590	13.2650	0.6300	21.1074
PDHKF				6.6640	0.5483	16.4820	0.5890	24.2833
TDHKF				6.6640	0.5483	32.2090	0.6090	40.0303
HKF	58%	75%	3 I_2	7.1338	0.5390	14.7850	0.8460	23.3038
PDHKF				7.1592	0.4908	18.3150	1.0450	27.0100
TDHKF				7.1592	0.4908	32.5320	1.3480	41.5300

The structure we proposed is suitable to be used in parallel and distributed processing environments and is applicable to multisensor networks, as it provides a significant reduction on the effects caused by communication and synchronization delays, as well as, it can avoid a generalized degradation of the estimation process, in opposition to what happens to the coordinated schemes, for example, when a fault in the central processor occurs.

Some important theoretical aspects for further investigation are briefly discussed in the following:

Suboptimality Evaluation

In the example in section 5, figures of merit were defined and used to compare the performances of the PDHKF and the HKF. As alternative figures of merit we propose the use of the traces of the covariance matrices in each iteration k as follows:

$$J(HKF)_k = \text{trace}(G\, P_k)$$

$$J(PDHKF)_k = \text{trace}\left(\sum_{i=1}^{N} G_i\, P_k^i\right)$$

where G and G_i are positive definite matrices, P is the error covariance matrix of the HKF, P^i the error covariance matrix of the PDHKF and N the number of subsystems.

In addition, we might define the PDHKF filter suboptimality degree at instant k by:

$$SD_k = \frac{J(HKF)_k - J(PDHKF)_k}{J(HKF)_k}$$

By evaluating this suboptimality degree, an objective assessment of the PDHKF filter can be carried out.

Stability of interconnected systems

For the heuristic partitioning method, into N subsystems, presented in this article, the i-th subsystem state is given by:

$$X^i_{k+1} = A^i_{k+1} X^i_k + \sum_{\substack{i=1 \\ i \neq j}}^{N} A^{ij}_{k+1} X^j_k + w^i_k$$

Even for the weak interaction partitioning criterion where the term $\sum_{\substack{i=1 \\ i\neq j}}^{N} A^{ij}_{k+1} X^j_k$ is relatively small, but has an influence on X^i_{k+1}, its consideration will lead to smaller negative effects on the quality of the estimation results than if we treated the global problem as a totally decoupled system with: $X^i_{k+1} = A^i_{k+1} X^i_k + w^i_k$, despite the quantitatively similar behavior between the totally decoupled system and the composite system with interaction terms in this case.

The resulting totally decoupled system might have an unstable subsystem while the original composite system is stable(Siljak, 1991).

For the general partitioning case via linear transformations:

$$X^j_k = T^j_k X_k$$

where T^j is the system partitioning matrix (n_j x n) applied on the global state for the j-th subsystem; it is assumed that T_j is of full rank n_j , $\forall$ j. This implies that the rank of the global system linear transformation matrix T is n , such that

$$X_k = \sum_{j=1}^{N} \Gamma^j_k X^j_k$$

where

Γ^j_k is the pseudo inverse matrix of T^j_k .

In this partitioning method, the transformation matrices T^j have to be chosen such that the resultant subsystems are stable (Willsky *et alii*,1982).

Interaction between the subsystems

Further investigation on the suboptimality degree is desirable in order to determine the criteria for evaluation of the proposed procedure. A possible direction for reaching this goal might be the utilization of the theory of perturbations (Hassan *et alii*, 1980), to evaluate error bounds in the PDHKF approach for various levels of interaction.

In this paper some figures of merit were proposed to somehow quantify suboptimality.

In spite of the suboptimality of the proposed approach, the results in Table 1 show very good state estimations.

Through the proposed figures of merit and suboptimality degree criteria, quantitative evaluations of our approach show its quality when applied to the example. From comparisons between the proposed structure and the one proposed in (Hassan *et alii*, 1978). , Table 1 and Fig.2 show very close estimation results.

Due to the fact of deriving a new estimation structure, based on an application of an estimation structure that had previously been developed, we were able : i) To conclude that in a decentralization procedure, a compromise between the suboptimality of the developed structure and its computational benefits on the distributed environment in use must be considered; ii) To acquire new insights about the design of distributed Kalman filtering algorithms, for example: Which structure can be derived from the application of the approach presented in this article to the hierarchical structure developed in (Hashemipour, Roy, and Laub, 1987). , based on information covariance matrix ?

REFERENCES

Hashemipour, R.H., Roy S., and Laub, J.A., (1987). Decentralized Structures for Kalman Filtering, IEEE Trans. On Automatic Control, vol.33,pp.88-94.

Hassan, M., Titli, A. Singh, M. G., and Hurteau, R., (1980), Stability and performance of large-scale hierarchical systems under certain structural perturbations, IEE Proc. D, Control and Theory Appl., No. 127, Vol.5, pp. 207-213.

Hassan, M.F., Salut, G., Singh, M.G,, and Titli A., (1978). A Decentralized Computational Algorithm For The Global Kalman Filter, IEEE Trans. On Automatic Control,vol.23,pp.262-267.

Kalman, R. E. (1960), A new approach to linear filtering and prediction problems, Trans. ASME, J. Basic Engng., vol.82, pp.34-45.

Quirino, R. B., Bottura, C.P., Costa Filho J. T., (1998). A Computational Structure For Parallel and Distributed Kalman Filtering., Proc. 12th Brazilian Automatic Control Conference, Uberlandia-MG, Brazil, 14-18 de Setembro de 1998.

Shah, M. M., (1971). Suboptimal Filtering Theory for Interacting Control Systems, Ph.D. Thesis, University of Cambridge.

Sherman, S., (1958). Non-mean-square error criteria, I.R.E Trans. On Information Theory, IT-4, pp. 125-126.

Siljak, D. D., (1991), Decentralized Control of Complex Systems, Mathematics In Science and Engineering, Vol. 184, Academic Press, pp. 65-118.

Tamura, H., (1974). A Discrete Model with Distributed Transport Delays and its Hierarchical Optimization for Preserving Stream Quality,. IEEE Trans. On Systems, Man, and Cybernetics,vol.SMC 4,no.5,p. 424.

Willsky, Alan S., Bello, Martin G., Castanon, David A., Levy, Bernard C., Verghese, George C., (1982). Combining and Updating of Local Estimates and Regional Maps Along Sets of One-Dimensional Tracks, IEEE Trans. On Automatic Control,vol.27,Nr.4,pp.799-812.

APPENDIX A

Approximate System and Measurement Disturbances as White Noise Sequences.

In section I of this appendix an explanation of why this is a reasonable approximation is heuristically suggested, and in section II the covariances of these disturbances are derived.

Section I

For simplicity the continuos time case is treated here.
The system and the exact Kalman filter can be represented in block diagram form in Fig. A1. Note that time arguments of x(t), u(t), etc, are ignored and these variables are written as x, u , etc.
A well known property of the Kalman filter is that the innovation signal $y - H\hat{x}$ is a white noise, say e:

$$y - H\hat{x} = e,$$

i.e., $H\tilde{x} = e - v$
This result states that certain linear combinations of the error signal $\tilde{x}$ are equal to linear combinations of white noises. Consider the case of most practical interest, i.e., relatively weak measurement noise v. When the measurement noise spectral density R is relatively small, the gain G in the loop constituting the filter (Fig. A1) is large (so that the filter output follows the filter input relatively close). Thus the bandwidth of the Kalman filter is relatively high. The high frequency components of its input y, which are due to v, then become less strongly attenuated. So that $\hat{x}$ contains relatively high frequency components. These components, being absent in x, contribute to the error $\tilde{x} = x - \hat{x}$. Plausibly, then the error $\tilde{x}$ contains high frequency components and can be approximated as white noise.
The following simple scalar example supports the above heuristic reasoning.
Fig. A2 shows that

$$\tilde{x} = x - \hat{x}$$

$$\tilde{x} = \frac{u}{G+1+s} - \frac{Gv}{G+1+s} \qquad \text{(A1)}$$

Hence the spectral density of $\tilde{x}$ is

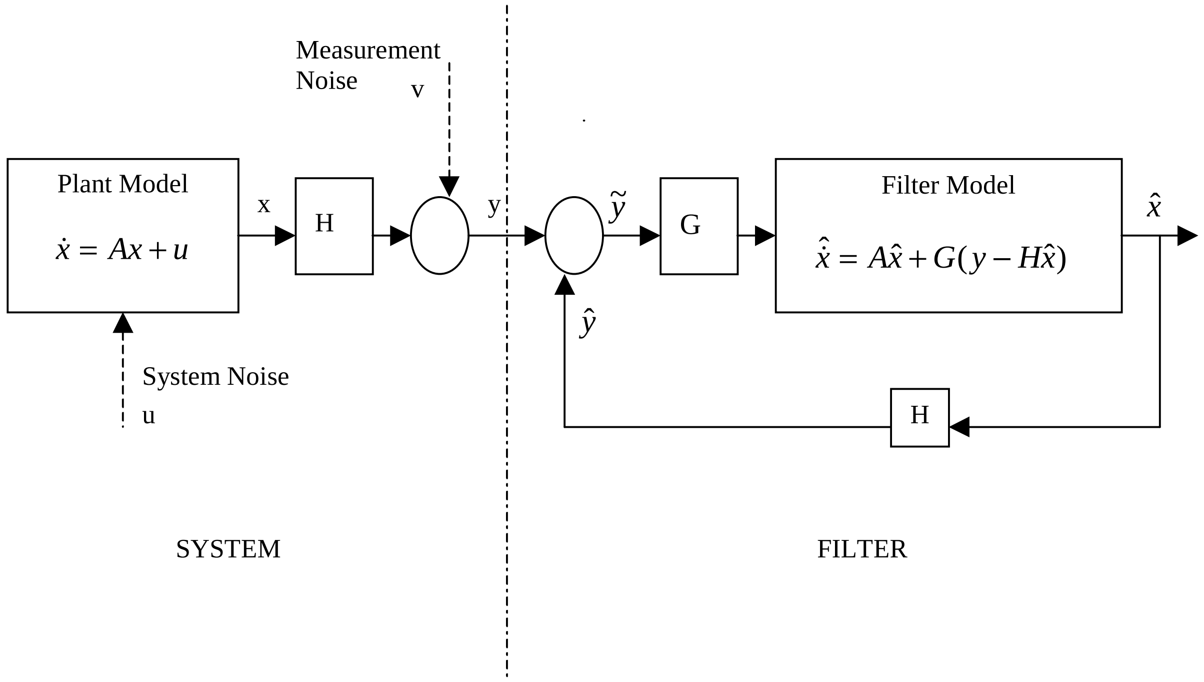

Fig.A1- System and Kalman-Bucy Filter

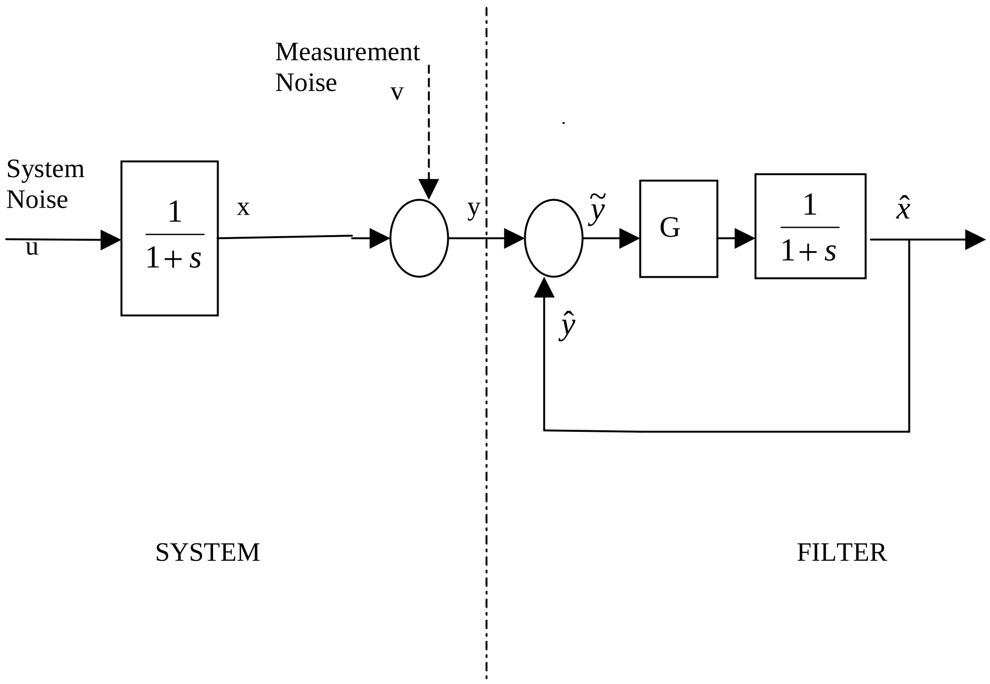

Fig.A2- Example: System and Filter

$$\Phi_{\tilde{x}}(w) = \frac{Q + G^2 R}{(G+1)^2 + w^2}, \qquad \text{(A2)}$$

where Q and R are the spectral densities of u and v. Similarly,

$$\Phi_{\tilde{y}}(w) = \frac{\left(\frac{Q}{R} + 1\right) + w^2}{(G+1)^2 + w^2} . R \qquad \text{(A3)}$$

Since $\tilde{y}$ is white , this spectral density is a constant. Hence, (A3) implies

$$G = \sqrt{\frac{Q}{R} + 1} - 1 \qquad \text{(A4)}$$

This result shows that when the measurement noise spectral density R is small, the filter gain G is indeed large.

In this case (A2) indicates that

$$\Phi_{\tilde{x}}(w) = \frac{R}{1+\frac{w^2}{G^2}} \cong R \quad \text{if } G \to \infty, \qquad \text{(A5)}$$

i.e., $\tilde{x}$ is approximately white.

On the other hand, when Q/R is small, i.e., when G is small, (A2) indicates that

$$\Phi_{\tilde{x}}(w) = \frac{Q}{1+w^2} = \Phi_x(w), \qquad \text{(A6)}$$

i.e., the spectral density of the error $\tilde{x}$ is approximately the same as that of the signal x. In this case relating $\tilde{x}$ by a white noise in the approximate system and measurement disturbances method is not a good approach. However, it may be a better approximation than neglecting $\tilde{x}$ altogether (see example – Section 5).

Section II

The approximate measurement and system disturbances for the two partitioned subsystems can be written as:

$$v_k^{*1} = v_k^1 + H_k^{12}\tilde{x}^2_{k/k-1}$$
$$v_k^{*2} = v_k^2 + H_k^{21}\tilde{x}^1_{k/k-1} \qquad \text{(A7)}$$

$$u_k^{*1} = u_k^1 + A_k^{12}\tilde{x}^2_{k/k}$$
$$u_k^{*2} = u_k^2 + A_k^{21}\tilde{x}^1_{k/k} \qquad \text{(A8)}$$

where

v_1, v_2, u_1 and u_2 are all uncorrelated zero mean white noise sequences and the remaining *a priori* and *a posteriori* estimation error terms have a Gaussian distribution (Kalman, 1960).

In order to derive the covariances of the approximate disturbances we consider the concept of innovation processes (Kalman, 1960).

Consider subsystem I: the new measurement information brought into subsystem I at time k is

$$e^1_{k/k-1} = H_k^{11}\tilde{x}^1_{k/k-1} + H_k^{12}\tilde{x}^2_{k/k-1} + v_k^1 \qquad \text{(A9)}$$

where $e^1_{k/k-1}$ is the white innovations sequence of covariance:

$$\begin{bmatrix} H_k^{11} & H_k^{12} \end{bmatrix} P^{11}_{k/k-1} \begin{bmatrix} H_k^{11} \\ H_k^{12} \end{bmatrix} + R_k^{11} \qquad \text{(A10)}$$

Now consider the covariance of the "partial" innovations:

$$v_k^{*1} = v_k^1 + H_k^{12}\tilde{x}^2_{k/k-1} \qquad \text{(A11)}$$

that is

$$\mathrm{E}(v_k^{*1} v_j^{*1^t}) = H_k^{12}\mathrm{E}(\tilde{x}^2_{k/k-1}\tilde{x}^{2^t}_{j/j-1})H_j^{12^t} + \mathrm{E}(v_k^1 v_j^{1^t})$$
$$+ H_k^{12}\mathrm{E}(\tilde{x}^2_{k/k-1}v_j^{1^t}) + \mathrm{E}(v_k^1\tilde{x}^{2^t}_{j/j-1}) \cdot H_j^{12^t} \qquad \text{(A12)}$$

for which there are three cases:

Case k>j

The second term is zero since v^1 is a white noise sequence. The fourth term is zero since future noise is independent of past errors in the state estimates.

Hence, combining the first and third terms we have

$$H_k^{12}\mathrm{E}[\tilde{x}^2_{k/k-1}(H_j^{12}\tilde{x}^2_{j/j-1} + v_j^1)^t] =$$
$$= H_k^{12}\mathrm{E}[\tilde{x}^2_{k/k-1}(y_j^{1^t} - \tilde{x}^{1^t}_{j/j-1}H_j^{11^t})] \qquad \text{(A13)}$$

By using Sherman's theorem (Sherman, 1958) and the theorem of orthogonal projections (Kalman, 1960), it follows that the first term is zero since the error at future time is orthogonal to past measurements. The second term in the above expression represents the correlation of the estimation errors for subsystems I and II. If the partitioning criterion is chosen such that the correlation between the subsystems is weak, then as a first approximation it may be possible to ignore such terms.

Hence, equation (A12) becomes

$$\mathrm{E}(v_k^{*1} v_j^{*t}) = \mathrm{E}(v_k^1 v_j^{1^t}) = 0$$

Similarly, we can show this to be true for the case k<j.

Case k=j

The third and fourth terms of equation (A12) are zero since the present subsystem I measurement noise is uncorrelated with the a-priori state estimate of subsystem II.

Hence,

$$\mathrm{E}(v_k^{*1} v_j^{*1^t}) = H_k^{12}P^{22}_{k/k-1}H_k^{12^t} + R_k^{11} \quad \text{for k=j}$$
$$= 0 \quad \text{for } k \neq j$$

Similarly, for subsystem II we can show that,

$$\mathrm{E}(v_k^{*2} v_j^{*2^t}) = H_k^{21}P^{11}_{k/k-1}H_k^{21^t} + R_k^{22} \quad \text{for k=j}$$
$$= 0 \quad \text{for } k \neq j$$

For the covariance of the approximate system disturbances u_k^{*1} and u_k^{*2} we also have three cases:

For instance, for subsystem I:

$$\mathrm{E}(u_k^{*1} u_j^{*1^t}) = \mathrm{E}[(A_k^{12}\tilde{x}^2_{k/k} + u_k^1)(A_j^{12}\tilde{x}^2_{j/j} + u_j^1)^t]$$
$$= \mathrm{E}[A_k^{12}\tilde{x}^2_{k/k}\tilde{x}^{2^t}_{j/j}A_j^{12^t} + u_k^1 u_j^{1^t}$$
$$+ A_k^{12}\tilde{x}^2_{k/k}u_j^{1^t} + u_k^1\tilde{x}^{2^t}_{j/j}A_j^{12^t}] \qquad \text{(A14)}$$

Case k>j

The second term of equation (A14) is zero since u^1 is a white noise sequence.

The fourth term is zero since the future subsystem I noise u^1 is uncorrelated with the estimation error $\tilde{x}^2_{j/j}$ of subsystem II.

Hence,

$$E(u_k^{*1} u_j^{*1^t}) = E[A_k^{12}\tilde{x}_{k/k}^2 \tilde{x}_{j/j}^{2^t} A_j^{12^t} + A_k^{12}\tilde{x}_{k/k}^2 u_j^{1^t}]$$

Case k<j (A15)

In this case the second and third terms in equation (A14) are zero.

Hence,

$$E(u_k^{*1} u_j^{*1^t}) = E[A_k^{12}\tilde{x}_{k/k}^2 \tilde{x}_{j/j}^{2^t} A_j^{12^t} + u_k^1 \tilde{x}_{j/j}^{2^t} A_j^{12^t}] \quad (A16)$$

Case k=j

The third and fourth terms in equation (A14) are zero since the present subsystem I noise u^1 is uncorrelated with the *a posteriori* estimation error $\tilde{x}_{k/k}^2$ of subsystem II.

Hence,

$$E(u_k^{*1} u_j^{*1^t}) = A_k^{12} P_{k/k}^{22} A_j^{12^t} + Q_k^{11} \quad (A17)$$

Now, u^{*1} can be approximated to be a white noise sequence if the expressions in the right hand side of equation (A15) and (A16) are insignificant compared to the terms in equation (A17).

For simplicity this aspect is illustrated by considering the continuos time case.

Consider the system and the filter equations,

$$\dot{x} = Ax + u$$

$$y = Hx + v$$

$$\dot{\hat{x}} = A\hat{x} + G(y - H\hat{x})$$

$$\dot{\tilde{x}} = (A - GH)\tilde{x} + u - Gv$$

Integrating the above expression between time instants j and k,

$$\tilde{x}_j = \exp((A - GH)(j - k))\tilde{x}_k$$

$$+ \int(\exp((A - GH)(j - \sigma)))(-G(\sigma)v(\sigma) + u(\sigma))d\sigma$$

Using the above equation,

$$E(\tilde{x}_j \tilde{x}_k^t)) = (\exp(A - GH)(j - k))E(x_k x_k^{\ t})$$

(A18)

$$E(\tilde{x}_j u_k^t) = (\exp(A - GH)(j - k))E(u_k u_k^t)$$

The terms on the right hand side of equation (A18) will be small in the practically important case of large gain G, i.e., case of small measurement noise levels.

Hence, for this class of systems the approximate system disturbance u_k^{*1} can be approximated to be a white noise sequence. Here the expression on the right hand side equations (A15) and (A16) are much smaller than the term in equation (A17).

In a similar manner for subsystem II, u_k^{*2} is a white noise sequence of covariance $(A_k^{21} P_{k/k}^{11} A_k^{21^t} + Q_k^{22})$.

Here the case of small measurement noise levels has been considered but in a practical study (see example–Section 5) : the **PDHKF** (Partially Decoupled Hierarchical Kalman Filter) filter leads to good results even for large measurement noise levels. This aspect needs more theoretical investigation.