



**HAL**  
open science

## Traitement de données contradictoires issues de systèmes multicapteurs

José Ragot, Didier Maquin

► **To cite this version:**

José Ragot, Didier Maquin. Traitement de données contradictoires issues de systèmes multicapteurs. Colloque TEC'88, Sep 1988, Grenoble, France. hal-00959708

**HAL Id: hal-00959708**

**<https://hal.science/hal-00959708>**

Submitted on 15 Mar 2014

**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.

José RAGOT, Didier MAQUIN

## **TRAITEMENT DE DONNEES CONTRADICTOIRES**

### **ISSUES DE SYSTEMES MULTICAPTEURS**

CRAN - URA DO 349 et CRVM UA 235  
Laboratoire d'Automatique et de Recherche Appliquée  
Rue du Doyen Marcel Roubault - B.P. 40  
54 501 Vandoeuvre les Nancy

#### *Résumé :*

*cette communication propose une méthodologie du traitement hiérarchisé de la validation de données contradictoires par équilibrage de bilan. Par suite des erreurs de mesure dues aux capteurs les mesures effectuées sur un procédé physique sont en général bruitées; le problème posé consiste alors à estimer les erreurs commises où à estimer les grandeurs vraies. Dans le cas de systèmes multicapteurs ce traitement peut être très complexe en particulier quand les modèles du procédé ne sont pas linéaires et quand les informations sont nombreuses.*

#### **1 - PRESENTATION DU PROBLEME**

La plupart des procédés peuvent être décrits par des équations de conservation de masse (ou d'énergie). Malheureusement les informations collectées sur ce procédé sont souvent contradictoires en ce sens qu'elles ne vérifient pas nécessairement ces équations de bilan; il est alors nécessaire de procéder à une réconciliation de ces données.

Dans cette étude on s'intéresse au cas des procédés en régime statique décrits par des équations multilinéaires; les erreurs de mesure sont supposées être des réalisations de variables aléatoires distribuées suivant des lois normales centrées (dans le cas de distribution non centrée la valeur moyenne représentant le biais du capteur de mesure devrait être estimé en tant que paramètre supplémentaire).

#### **2 - FORMULATION DU PROBLEME**

La classe des systèmes multilinéaires correspond physiquement à la structure des modèles de connaissance pour des systèmes physiques en régime permanent décrits par des équations de bilan en flux total et en flux partiel. Les modèles se représentent sous forme structurale exacte par un ensemble d'équations du type :

$$f_j = M \prod_{i=1}^j \hat{X}_i^* \quad j= 1, \dots, d \quad (1)$$

où  $d$  représente l'ordre maximal du modèle,  
 $M$  la matrice d'incidence du réseau du procédé,  
 $\hat{X}_j^*$  le vecteur des grandeurs réelles relatives à la  $j$ ème variable,  
l'opérateur qui associe à deux vecteurs un troisième vecteur dont les composantes sont les produits des composantes de même rangs de ces deux vecteurs (dans la suite on, notera  $*$  ce produit).

A titre d'exemple :

$$f_1 = M \hat{Q}^* = 0 \quad (2)$$

$$f_2 = M \hat{Q}^* \hat{d}^* = 0 \quad (3)$$

pourraient représenter les équations de conservation des flux totaux volumique ( $Q_j$ ) et les conservations en flux totaux massiques ( $Q_i d_i$ ). Pratiquement on a accès à des formes quadrilinéaires au maximum quand on prend en compte des débits volumiques, des masses volumiques, des teneurs et des granulométries.

Le problème de la validation de données par équilibrage de bilan consiste à déterminer les estimations  $\hat{X}$  des grandeurs vraies  $\hat{X}^*$ , à partir des mesures  $X$ , qui minimisent le critère  $J$  :

$$J = \frac{1}{2} \sum_{j=1}^d (\hat{X}_j - X_j)^T V_j (\hat{X}_j - X_j) \quad (4)$$

où  $V_j$  est la matrice de variance-covariance des erreurs de mesure.

A ce problème on associe le Lagrangien  $L$  :

$$L = \frac{1}{2} \sum_{j=1}^d [(\hat{X}_j - X_j)^T V_j (\hat{X}_j - X_j) + \lambda_j^T M \prod_{i=1}^j \hat{X}_i] \quad (5)$$

où  $\lambda_j$  est le vecteur des paramètres de Lagrange associé à la  $j$ ème équation de bilan.

L'estimateur cherché rend stationnaire le Lagrangien, ce qui s'explique :

$$V_j (\hat{X}_j - X_j) + \sum_{k=1}^d (M \prod_{i=1, i \neq j}^k \hat{X}_i)^T \lambda_k = 0 \quad (6)$$

$$M \prod_{i=1}^j \hat{X}_i = 0 \quad (7)$$

Ce système est de très grande dimension et non linéaire; cependant, sa formulation souligne l'intérêt d'une résolution par décomposition hiérarchisée des calculs.

### 3 - RESOLUTION HIERARCHISEE

Afin de mettre en évidence plus clairement le principe de cette solution on raisonne sur le cas de modèles quadrilinéaires ce qui suffit pour entrevoir la généralisation à des systèmes n-linéaires.

Le Lagrangien s'explique (avec des variances unités pour simplifier les écritures) :

$$L = \frac{1}{2} \sum_{j=1}^4 [ (\hat{X}_j - X_j)^T V_j (\hat{X}_j - X_j) + \lambda_j^T M \prod_{i=1}^j \hat{X}_i ] \quad (8)$$

Dans la suite on allège les notations en posant :

$$M * \hat{X}_i * \hat{X}_j = M_{ij} \quad (9)$$

Les conditions de stationnarité du Lagrangien s'explicitent :

$$\hat{X}_1 - X_1 + M^T \lambda_1 + M^T_2 \lambda_2 + M^T_{23} \lambda_3 + M^T_{234} \lambda_4 = 0 \quad (10)$$

$$M \hat{X}_1 = 0 \quad (11)$$

$$\hat{X}_2 - X_2 + M^T_1 \lambda_2 + M^T_{13} \lambda_3 + M^T_{134} \lambda_4 = 0 \quad (12)$$

$$M_1 \hat{X}_2 = 0 \quad (13)$$

$$\hat{X}_3 - X_3 + M^T_{12} \lambda_3 + M^T_{124} \lambda_4 = 0 \quad (14)$$

$$M_{12} \hat{X}_3 = 0 \quad (15)$$

$$\hat{X}_4 - X_4 + M^T_{123} \lambda_4 = 0 \quad (16)$$

$$M_{123} \hat{X}_4 = 0 \quad (17)$$

La résolution de ce système fortement non linéaire peut être réalisée par niveau ; le niveau d'indice i est relatif à l'estimation de  $X_i$ .

Prémultiplions l'équation (10) par  $M^T$ ; on extrait alors la valeur du paramètre de Lagrange dont le report dans (10) explicite alors l'estimé  $X_1$  sous la forme :

$$\hat{X}_1 = [I - M^T (MM^T)^{-1} M] [X_1 - M^T_2 \lambda_2 - M^T_{23} \lambda_3 - M^T_{234} \lambda_4] \quad (18)$$

Pour simplifier cette écriture définissons la matrice diagonale  $\Lambda_i$  dont les éléments

diagonaux sont égaux aux composantes de  $X_i$ . Alors l'expression (9) peut s'écrire aussi :

$$M_{ij} = M \Lambda_i \Lambda_j \quad (19)$$

et l'estimateur  $\hat{X}_1$

:

$$\hat{X}_1 = [I - M^T (MM^T)^{-1} M] [X_1 - \Lambda_2 (M^T \lambda_2 + \Lambda_3 M^T \lambda_3 + \Lambda_3 \Lambda_4 M^T \lambda_4)] \quad (20)$$

De même, avec la notation (19), l'équation (12) s'écrit :

$$\hat{X}_2 - X_2 + \Lambda_1 (M^T \lambda_2 + \Lambda_3 M^T \lambda_3 + \Lambda_3 \Lambda_4 M^T \lambda_4) = 0 \quad (21)$$

$$\hat{X}_1 = [I - M^T (MM^T)^{-1} M] [X_1 - \Lambda_2 \Lambda^{-1}_1 (\hat{X}_2 - X_2)] \quad (22)$$

De façon analogue on montre que :

$$\hat{X}_2 = [I - M^T_1 (M_1 M^T_1)^{-1} M_1] [X_2 - \Lambda_3 \Lambda^{-1}_2 (\hat{X}_3 - X_3)] \quad (23)$$

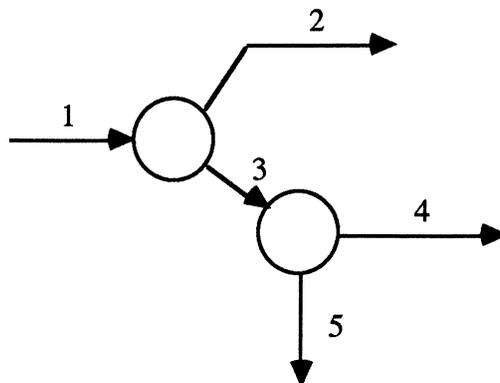
$$\hat{X}_3 = [I - M^T_{12} (M_{12} M^T_{12})^{-1} M_{12}] [X_3 - \Lambda_4 \Lambda^{-1}_3 (\hat{X}_4 - X_4)] \quad (24)$$

$$\hat{X}_4 = [I - M^T_{123} (M_{123} M^T_{123})^{-1} M_{123}] [X_4] \quad (25)$$

Ainsi chaque niveau de calcul met en oeuvre une procédure générale de projection des mesures (modifiées par les estimations des niveaux postérieurs) sur l'espace des contraintes de bilan. Les équations (21) à (25) se résolvent directement en choisissant comme estimations initiales les mesures; le calcul se présente sous forme d'une série d'itérations directes, dont la convergence peut toujours être assurée en appliquant un coefficient de relaxation à chaque estimation.

#### 4 - EXEMPLE NUMERIQUE

On considère un exemple simple en minéralurgie à propos d'une cellule de flottation chimique schématisée par deux noeuds (cellules) et cinq arcs (voies de circulation de la matière).



Sur chaque voie on dispose de mesure du débit volumique  $Q_i$ , de masse volumique  $\rho_i$ , de la proportion  $t_i$  de particules appartenant à la classe granulométrique  $i$  et enfin de la concentration chimique  $c_i$  d'une espèce minérale appartenant à cette classe granulométrique.

Le tableau 1 regroupe les mesures effectuées sur ces différentes variables.

Pour visualiser les incohérences des mesures brutes, on a calculé les résidus de bilan pour ces deux noeuds :

$$\begin{aligned}
 r_{11} &= Q_1 - Q_2 - Q_3 &= -1 \\
 r_{12} &= Q_3 - Q_4 - Q_5 &= 1 \\
 r_{21} &= Q_1 \rho_1 - Q_2 \rho_2 - Q_3 \rho_3 &= -2 \\
 r_{22} &= Q_3 \rho_3 - Q_4 \rho_4 - Q_5 \rho_5 &= 1.6 \\
 r_{31} &= Q_1 \rho_1 t_1 - Q_2 \rho_2 t_2 - Q_3 \rho_3 t_3 &= -1.428 \\
 r_{32} &= Q_3 \rho_3 t_3 - Q_4 \rho_4 t_4 - Q_5 \rho_5 t_5 &= 2.116 \\
 r_{41} &= Q_1 \rho_1 t_1 c_1 - Q_2 \rho_2 t_2 c_1 - Q_3 \rho_3 t_3 c_1 &= 2.213 \\
 r_{42} &= Q_3 \rho_3 t_3 c_3 - Q_4 \rho_4 t_4 c_4 - Q_5 \rho_5 t_5 c_5 &= -.114
 \end{aligned}$$

La procédure exposée précédemment a été appliquée pour réconcilier l'ensemble des données. Le tableau 2 donne les estimations obtenues et qui vérifient les contraintes de bilan.

| Voie | Débit volumique | Masse volumique | Distribution granulométrique | Concentration |
|------|-----------------|-----------------|------------------------------|---------------|
| 1    | 10.000          | 1.500           | 70.00                        | 1.200         |
| 2    | 2.000           | 1.300           | 60.00                        | 4.000         |
| 3    | 9.000           | 1.600           | 72.00                        | 0.400         |
| 4    | 4.000           | 1.500           | 65.00                        | 0.200         |
| 5    | 4.000           | 1.700           | 64.00                        | 0.800         |

Tableau 1 : mesures brutes

| Voie | Débit volumique | Masse volumique | Distribution granulométrique | Concentration |
|------|-----------------|-----------------|------------------------------|---------------|
| 1    | 10.261          | 1.536           | 65.73                        | 1.0711        |
| 2    | 1.773           | 1.339           | 69.36                        | 4.0205        |
| 3    | 8.487           | 1.577           | 65.09                        | 0.5136        |
| 4    | 4.224           | 1.477           | 65.71                        | 0.1975        |
| 5    | 4.243           | 1.676           | 64.55                        | 0.7973        |

Tableau 2 : données réconciliées

## 5 - CONCLUSION

L'algorithme à itération directe que nous proposons présente différents attraits :

- il résoud le problème d'équilibrage de bilan de façon optimale,
- il s'applique bien aux problèmes de très grande dimension,
- il met en oeuvre des calculs simples à chaque niveau,
- la prise en compte des interactions est faite sans variables supplémentaires.

Pour être complet, il resterait à expliciter les caractéristiques statistiques de cet estimateur et en particulier sa sensibilité vis à vis des erreurs de mesure ainsi que sa variance; ce calcul ne peut être fait qu'en linéarisant les différentes équations de contraintes autour de la solution obtenue après équilibrage.

Enfin dans un cas plus général il serait intéressant d'examiner le problème de l'observabilité; actuellement nous disposons d'une solution dans le cas bilinéaire mais il reste à étendre ce résultat dans le cas n-linéaire.