



HAL
open science

Méthodes de compensation de biais pour l'identification de modèles erreurs-en-les-variables

Stéphane Thil, Marion Gilson, Hugues Garnier

► **To cite this version:**

Stéphane Thil, Marion Gilson, Hugues Garnier. Méthodes de compensation de biais pour l'identification de modèles erreurs-en-les-variables. Journées Doctorales / Journées Nationales MACS, Sep 2005, Lyon, France. pp.CD-ROM. <hal-00152700>

HAL Id: hal-00152700

<https://hal.science/hal-00152700v1>

Submitted on 7 Jun 2007

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.



HAL Authorization

MÉTHODES DE COMPENSATION DE BIAIS POUR L'IDENTIFICATION DE MODÈLES ERREURS-EN-LES-VARIABLES

Stéphane Thil, Marion Gilson et Hugues Garnier

Centre de Recherche en Automatique de Nancy - CRAN UMR 7039 CNRS-UHP-INPL,
Université Henri Poincaré, Nancy 1, BP 239 - 54506 Vandoeuvre Cedex
{stephane.thil, marion.gilson, hugues.garnier}@cran.uhp-nancy.fr

Résumé

Cet article traite de l'identification de modèles erreurs-en-les-variables, où les mesures de l'entrée et de la sortie sont contaminées par un bruit additif. Parmi les diverses méthodes existantes, nous nous intéressons ici à la famille de méthodes dites de compensation de biais. Trois méthodes récemment développées sont analysées et comparées à l'aide de simulations de Monte Carlo.

Mots clés : Identification des systèmes, erreurs-en-les-variables, compensation de biais.

Abstract

This paper deals with identification of errors-in-variables models, where both input and output measurements are corrupted by additive noise. Among the various existing techniques, we focus on the bias-eliminating methods. Three recently developed methods are exposed and compared with the help of Monte Carlo simulations.

Keywords : System identification, errors-in-variables, bias compensation.

INTRODUCTION

L'élaboration de modèles de systèmes dynamiques représente un point clé en automatique et dans les sciences expérimentales. L'identification consiste à rechercher un modèle d'un système à partir de données expérimentales et de connaissances disponibles *a priori*. Ce modèle doit être capable de reproduire le comportement dynamique du système réel de façon acceptable.

Ce domaine de recherche a fait l'objet de nombreuses études durant les trente dernières années. La plupart des méthodes d'identification développées reposent sur l'hypothèse de bruit additif sur la sortie uniquement [4, 5, 8]. Cependant, dans de nombreux cas pratiques, l'entrée est également bruitée, par exemple à cause de la présence d'erreurs de mesure. Or il s'avère que l'utilisation des méthodes classiques ne prenant pas en compte les bruits sur l'entrée conduit à des estimations biaisées. C'est pourquoi le problème pour lequel les données d'entrée/sortie sont observées à une erreur additive près, généralement dénommé 'erreurs-en-les-variables', est reconnu comme étant beaucoup plus délicat à traiter.

Ce thème de recherche a depuis peu fait l'objet de nombreux développements au sein de la communauté automatique [7, 6]. Il existe plusieurs méthodes pour traiter ce problème, suivant les connaissances *a priori* disponibles et les hypothèses sur les bruits en entrée et en sortie. Parmi celles-ci, nous nous intéressons dans cet article à la famille de méthodes dites de compensation de biais développées sous l'impulsion de W.X. Zheng. Le principe de ces techniques consiste à estimer le biais, puis à le soustraire des paramètres estimés par moindres carrés. Plusieurs approches utilisant cette technique ont été récemment proposées, notamment

pour l'identification des systèmes bouclés [2, 3]. L'objectif de cet article est d'analyser, puis de comparer, trois méthodes de compensation de biais récemment développées (à savoir BELS1 [9], BELS1-II [10], et FBELS [11]) pour l'identification des modèles erreurs-en-les-variables avec bruits blancs en entrée et en sortie.

POSITION DU PROBLÈME

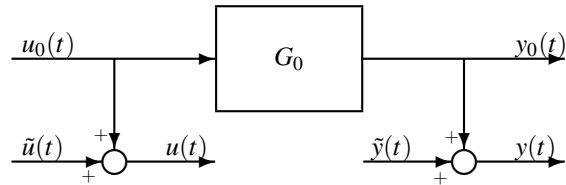


FIG. 1 – Système avec données d'entrée/sortie bruitées

Considérons le système linéaire invariant suivant décrit par la figure (1) :

$$\mathcal{S} : \begin{cases} y_0(t) = G_0(q)u_0(t) \\ u(t) = u_0(t) + \tilde{u}(t) \\ y(t) = y_0(t) + \tilde{y}(t) \end{cases} \quad (1)$$

où q désigne l'opérateur avance, et :

- $G_0(q) = B_0(q^{-1})/A_0(q^{-1})$: opérateur de transfert du vrai système ;
- $u_0(t)$: signal d'entrée non bruité ;
- $y_0(t)$: signal de sortie non bruité ;
- $\tilde{u}(t)$: bruit de mesure sur l'entrée ;
- $\tilde{y}(t)$: bruit de mesure sur la sortie ;
- $u(t)$: signal d'entrée mesuré ;
- $y(t)$: signal de sortie mesuré.

Le système ainsi décrit prend en compte un bruit additif en sortie ainsi qu'un bruit additif en entrée, et est connu sous le nom de modèle dynamique avec erreurs-en-les-variables.

L'application des méthodes considérées dans cet article nécessite de faire les hypothèses suivantes :

- H1. la fonction de transfert $G_0(z)$ est asymptotiquement stable ;
- H2. les polynômes $A_0(z)$ et $B_0(z)$ n'ont pas de facteurs communs ;
- H3. le signal d'entrée non bruité $u_0(t)$ est ergodique et stationnaire au sens large ;
- H4. $\tilde{u}(t)$ et $\tilde{y}(t)$ sont des bruits blancs de moyenne nulle et de variance respective $\lambda_{\tilde{u}}$ et $\lambda_{\tilde{y}}$;
- H5. $u_0(t)$, $\tilde{u}(t)$, $\tilde{y}(t)$ sont mutuellement décorrélés ;
- H6. les degrés n_a et n_b des polynômes $A_0(q^{-1})$ et $B_0(q^{-1})$ sont connus.

En particulier, comme $G_0(z)$ est stable et $u_0(t)$, $\tilde{u}(t)$ et $\tilde{y}(t)$ sont stationnaires au sens large, les signaux mesurés sont aussi stationnaires au sens large. Subséquemment, leur fonction d'autocorrélation ne dépend pas des instants t_1, t_2 auxquels elle est considérée, mais de la différence $t_1 - t_2$ ¹ :

$$r_{yy}(n) \triangleq \bar{E}[y(t)y(t-n)]; \quad r_{uy}(n) \triangleq \bar{E}[u(t)y(t-n)]$$

Introduisons le modèle paramétrique suivant :

$$\mathcal{M}(\theta) : \begin{cases} y_0(t) = G(q, \theta)u_0(t) \\ u(t) = u_0(t) + \tilde{u}(t) \\ y(t) = y_0(t) + \tilde{y}(t) \end{cases} \quad (2)$$

où l'on note :

$$\mathcal{G}(\theta) : G(q, \theta) = \frac{B(q^{-1}, \theta)}{A(q^{-1}, \theta)} = \frac{b_1q^{-1} + \dots + b_{n_b}q^{-n_b}}{1 + a_1q^{-1} + \dots + a_{n_a}q^{-n_a}}$$

$$\theta^T = [a^T \ b^T] = [a_1 \ \dots \ a_{n_a} \ b_1 \ \dots \ b_{n_b}] \in \mathbb{R}^{n_a+n_b}$$

Définissons de plus les vecteurs suivants :

$$\varphi^T(t) = \varphi_0^T(t) + \tilde{\varphi}^T(t) \in \mathbb{R}^{n_a+n_b} \quad (3)$$

$$\varphi_0^T(t) = [-y_0^T(t) \ u_0^T(t)]; \quad \tilde{\varphi}^T(t) = [-\tilde{y}^T(t) \ \tilde{u}^T(t)]$$

$$y_0^T(t) = [y_0(t-1) \ \dots \ y_0(t-n_a)]; \quad u_0^T(t) = [u_0(t-1) \ \dots \ u_0(t-n_b)]$$

$$\tilde{y}^T(t) = [\tilde{y}(t-1) \ \dots \ \tilde{y}(t-n_a)]; \quad \tilde{u}^T(t) = [\tilde{u}(t-1) \ \dots \ \tilde{u}(t-n_b)]$$

Alors le modèle (2) peut se réécrire sous la forme d'une régression linéaire :

$$y(t) = \varphi^T(t)\theta + \varepsilon(t) \quad (4)$$

où l'erreur d'équation $\varepsilon(t)$ s'écrit :

$$\varepsilon(t) = \tilde{y}(t) - \tilde{\varphi}^T(t)\theta \quad (5)$$

L'estimée asymptotique de θ au sens des moindres carrés est donnée par :

$$\theta_{LS}^* = \arg \min_{\theta} J(\theta) = \arg \min_{\theta} \bar{E}[\varepsilon^2(t)]$$

Si l'on note $R_{\varphi\varphi} \triangleq \bar{E}[\varphi(t)\varphi^T(t)]$ et $r_{\varphi y} \triangleq \bar{E}[\varphi(t)y(t)]$, la solution de cette minimisation s'écrit :

$$\theta_{LS}^* = R_{\varphi\varphi}^{-1}r_{\varphi y} = \theta - R_{\varphi\varphi}^{-1} \begin{pmatrix} \lambda_{\tilde{y}}I_{n_a} & 0 \\ 0 & \lambda_{\tilde{u}}I_{n_b} \end{pmatrix} \theta \quad (6)$$

où I_n représente la matrice identité de dimension n . Notons que pour obtenir (6) on a utilisé les équations (4), (5) et l'hypothèse H4.

On constate ainsi que la solution au sens des moindres carrés est asymptotiquement biaisée, et que le biais est entièrement déterminé par $\lambda_{\tilde{u}}$ et $\lambda_{\tilde{y}}$. Cela suggère donc une méthode de compensation de biais fondée sur l'estimation de ces deux variances.

¹ \bar{E} est l'espérance mathématique généralisée [5], définie par $\bar{E}[y(t)] \triangleq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{t=1}^N E[y(t)]$ où N représente le nombre de données disponibles.

MÉTHODE BELS1

Supposant que l'on dispose d'estimées des variances des bruits $\hat{\lambda}_{\tilde{u}}$ et $\hat{\lambda}_{\tilde{y}}$, il est possible d'obtenir une estimée non biaisée de θ par la procédure itérative suivante² [9] :

$$\hat{\theta}_{BELS}^{(i+1)}(N) = \hat{\theta}_{LS}(N) + \hat{R}_{\varphi\varphi}^{-1}(N) \begin{pmatrix} \hat{\lambda}_{\tilde{y}}I_{n_a} & 0 \\ 0 & \hat{\lambda}_{\tilde{u}}I_{n_b} \end{pmatrix} \hat{\theta}_{BELS}^{(i)}(N) \quad (7)$$

$$\text{avec } \hat{R}_{\varphi\varphi}(N) = \frac{1}{N} \sum_{t=1}^N \varphi(t)\varphi^T(t)$$

Il faut ainsi obtenir des estimées $\hat{\lambda}_{\tilde{u}}$ et $\hat{\lambda}_{\tilde{y}}$. Pour ce faire, W.X. Zheng propose d'augmenter artificiellement de un le degré du dénominateur de $G(q, \theta)$, avec le nouveau paramètre ainsi introduit ayant une valeur vraie égale à zéro. Définissons donc l'opérateur de transfert augmenté par :

$$\bar{G}(q, \theta) = \frac{B(q^{-1}, \theta)}{\bar{A}(q^{-1}, \theta)} = \frac{b_1q^{-1} + \dots + b_{n_b}q^{-n_b}}{1 + a_1q^{-1} + \dots + a_{n_a}q^{-n_a} + a_{n_a+1}q^{-n_a-1}}$$

et le vecteur des paramètres augmenté par :

$$\bar{\theta}^T = [a^T \ a_{n_a+1} \ b^T] \triangleq [\bar{a}^T \ b^T]$$

Bien que l'opérateur de transfert augmenté $\bar{G}(q, \theta)$ soit équivalent à l'opérateur de transfert original dans le sens où $a_{n_a+1} = 0$, l'augmentation du degré du dénominateur va permettre d'obtenir une équation linéaire que l'on pourra utiliser pour estimer les variances des bruits.

De la même manière que précédemment (voir équations (4) et (5)), on peut écrire le système augmenté sous la forme d'une régression linéaire :

$$y(t) = \bar{\varphi}^T(t)\bar{\theta} + \tilde{y}(t) - \tilde{\varphi}^T(t)\bar{\theta}$$

où :

$$\bar{\varphi}^T(t) = \bar{\varphi}_0^T(t) + \tilde{\varphi}^T(t) \quad (8)$$

$$\bar{\varphi}_0^T(t) = [-y_0^T(t) - y_0(t-n_a-1) \ u_0^T(t)]$$

$$\tilde{\varphi}^T(t) = [-\tilde{y}^T(t) - \tilde{y}(t-n_a-1) \ \tilde{u}^T(t)]$$

En comparant les définitions du vecteur de régression normal (3) et du vecteur de régression augmenté (8), on remarque qu'augmenter de un le degré du dénominateur de $G(q, \theta)$ revient en fait à utiliser une mesure supplémentaire de la sortie. L'estimateur des moindres carrés du vecteur augmenté $\bar{\theta}$ s'écrit :

$$\bar{\theta}_{LS}^* = R_{\bar{\varphi}\bar{\varphi}}^{-1}r_{\bar{\varphi}y} \quad (9)$$

où $R_{\bar{\varphi}\bar{\varphi}} \triangleq \bar{E}[\bar{\varphi}(t)\bar{\varphi}^T(t)]$ et $r_{\bar{\varphi}y} \triangleq \bar{E}[\bar{\varphi}(t)y(t)]$. Par analogie avec (6) on a :

$$\bar{\theta}_{LS}^* = \bar{\theta} - R_{\bar{\varphi}\bar{\varphi}}^{-1} \begin{pmatrix} \lambda_{\tilde{y}}I_{n_a+1} & 0 \\ 0 & \lambda_{\tilde{u}}I_{n_b} \end{pmatrix} \bar{\theta} \quad (10)$$

Comme $\bar{A}(q^{-1}, \theta)$ a un paramètre connu égal à zéro, le vecteur des paramètres $\bar{\theta}$ doit vérifier :

$$H^T \bar{\theta} = 0$$

²La convergence asymptotique de la procédure itérative est en cours d'étude et n'est donc pas démontrée dans cet article.

où $H^T = [h^T \ 0^T] \in \mathbb{R}^{n_a+n_b+1}$ et $h^T = [0 \ \dots \ 0 \ 1] \in \mathbb{R}^{n_a+1}$. En prémultipliant (10) par H^T , on obtient donc :

$$H^T \bar{\theta}_{LS}^* = -H^T R_{\bar{\phi}\bar{\phi}}^{-1} \begin{pmatrix} \lambda_{\bar{y}} I_{n_a+1} & 0 \\ 0 & \lambda_{\bar{u}} I_{n_b} \end{pmatrix} \bar{\theta} \quad (11)$$

Si l'on décompose la matrice $R_{\bar{\phi}\bar{\phi}}^{-1}$ en quatre sous-matrices de dimensions appropriées de la manière suivante :

$$R_{\bar{\phi}\bar{\phi}}^{-1} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$$

alors l'équation (11) se réécrit sous la forme :

$$-h^T \bar{a}_{LS}^* = \lambda_{\bar{y}} h^T M_{11} \bar{a} + \lambda_{\bar{u}} h^T M_{12} b \quad (12)$$

ce qui donne une première équation linéaire avec les deux inconnues à estimer, $\lambda_{\bar{y}}$ et $\lambda_{\bar{u}}$. La deuxième équation est obtenue en étudiant l'erreur moyenne de la méthode des moindres carrés, qui s'écrit [9] :

$$J(\bar{\theta}_{LS}^*) = \lambda_{\bar{y}} (1 + \bar{a}_{LS}^{*T} \bar{a}) + \lambda_{\bar{u}} b_{LS}^{*T} b \quad (13)$$

D'où l'algorithme BELS1 :

1. estimation du vecteur des paramètres augmenté $\bar{\theta}$ à partir des données $\{u(t), y(t) | t = 1, \dots, N\}$ disponibles par la méthode des moindres carrés, ce qui donne le vecteur des paramètres initial $\hat{\theta}_{BELS}^{(0)}(N) = \hat{\theta}_{LS}(N)$;
2. calcul de $\hat{\lambda}_{\bar{y}}^{(i+1)}$ et $\hat{\lambda}_{\bar{u}}^{(i+1)}$ en résolvant le système suivant, où les quantités asymptotiques ont été remplacées par leur estimation obtenue pour le nombre fini de données disponibles³ :

$$\begin{pmatrix} \hat{J}(\hat{\theta}_{LS}(N)) \\ -h^T \hat{a}_{LS}(N) \end{pmatrix} = \bar{\Delta} \begin{pmatrix} \hat{\lambda}_{\bar{y}}^{(i+1)} \\ \hat{\lambda}_{\bar{u}}^{(i+1)} \end{pmatrix} \quad (14)$$

avec :

$$\bar{\Delta} \triangleq \begin{pmatrix} 1 + \hat{a}_{LS}^T(N) \hat{a}_{BELS}^{(i)}(N) & \hat{b}_{LS}^T(N) \hat{b}_{BELS}^{(i)}(N) \\ h^T \hat{M}_{11} \hat{a}_{BELS}^{(i)}(N) & h^T \hat{M}_{12} \hat{b}_{BELS}^{(i)}(N) \end{pmatrix}$$

3. calcul de $\hat{\theta}_{BELS}^{(i+1)}(N)$:

$$\hat{\theta}_{BELS}^{(i+1)}(N) = \hat{\theta}_{LS}(N) + \hat{R}_{\bar{\phi}\bar{\phi}}^{-1}(N) \begin{pmatrix} \hat{\lambda}_{\bar{y}}^{(i+1)} I_{n_a+1} & 0 \\ 0 & \hat{\lambda}_{\bar{u}}^{(i+1)} I_{n_b} \end{pmatrix} \hat{\theta}_{BELS}^{(i)}(N)$$

4. répéter les étapes 2 et 3 jusqu'à convergence (selon le critère d'arrêt choisi).

Cette première méthode de compensation de biais dédiée aux modèles erreurs-en-les-variables permet d'obtenir des estimées asymptotiquement non biaisées des paramètres. Toutefois, dans le calcul des estimées de $\hat{\lambda}_{\bar{y}}$ et $\hat{\lambda}_{\bar{u}}$ par la résolution du système linéaire (14), elle utilise le vecteur des paramètres augmenté, ainsi que la matrice de covariance du vecteur de régression augmenté, induisant des calculs inutiles. Pour obvier à cet inconvénient, une nouvelle version a été proposée [10].

³Rappelons que $\frac{1}{N} \sum_{t=1}^N y(t)$ est un estimateur de $\bar{E}[y(t)]$.

MÉTHODE BELS1-II

L'objectif est donc ici d'exprimer le biais asymptotique sans faire de calculs où interviennent des variables augmentées, évitant ainsi des calculs superflus. Une première constatation est que l'équation de l'erreur moyenne asymptotique de la méthode des moindres carrés (13) est également vraie pour le vecteur des paramètres non augmenté :

$$J(\theta_{LS}^*) = \lambda_{\bar{y}} (1 + a_{LS}^{*T} a) + \lambda_{\bar{u}} b_{LS}^{*T} b \quad (15)$$

D'autre part, le coût algorithmique de la méthode BELS1 peut être réduit en remplaçant l'équation (12) par une autre, ne faisant pas intervenir de variables augmentées. L'idée est de construire le vecteur des paramètres augmenté en plaçant le nouveau paramètre a_{n+1} à la fin du vecteur θ . En effet, changeant les notations précédentes, si l'on construit le vecteur des paramètres augmentés de la manière suivante :

$$\bar{\theta}^T = [\theta^T \ a_{n_a+1}]$$

$$\text{et donc } \bar{\phi}^T(t) = [\phi^T(t) \ -y(t-n_a-1)]$$

il est alors possible d'exploiter la forme particulière du vecteur de régression $\bar{\phi}(t)$ pour décomposer le vecteur $r_{\bar{\phi}y}$ et la matrice $R_{\bar{\phi}\bar{\phi}}$ sous la forme :

$$r_{\bar{\phi}y} = \begin{pmatrix} r_{\phi y} \\ r_{yy}(n_a+1) \end{pmatrix} \quad (16)$$

$$R_{\bar{\phi}\bar{\phi}} = \begin{bmatrix} R_{\phi\phi} & \mathbf{v} \\ \mathbf{v}^T & r_{yy}(0) \end{bmatrix}$$

avec $\mathbf{v} = \bar{E}[\phi(t)y(t-n_a-1)]$. L'inverse de cette matrice s'écrit :

$$R_{\bar{\phi}\bar{\phi}}^{-1} = \begin{bmatrix} R_{\phi\phi}^{-1} + R_{\phi\phi}^{-1} \mathbf{v} \delta \mathbf{v}^T R_{\phi\phi}^{-1} & -R_{\phi\phi}^{-1} \mathbf{v} \delta^{-1} \\ -\delta^{-1} \mathbf{v}^T R_{\phi\phi}^{-1} & \delta^{-1} \end{bmatrix} \quad (17)$$

$$\text{avec } \delta = r_{yy}(0) - \mathbf{v}^T R_{\phi\phi}^{-1} \mathbf{v} \quad (18)$$

En injectant les équations (16), (17), (18) ainsi obtenues dans l'expression de l'estimateur des moindres carrés (9) on obtient :

$$(a_{n_a+1})_{LS}^* = \delta^{-1} r_{yy}(n_a+1) - \delta^{-1} \mathbf{v}^T R_{\phi\phi}^{-1} r_{\phi y} \quad (19)$$

Le changement de l'agencement du vecteur $\bar{\theta}$ implique que l'estimée au sens des moindres carrés (10) se réécrit :

$$\bar{\theta}_{LS}^* = \bar{\theta} - R_{\bar{\phi}\bar{\phi}}^{-1} \begin{pmatrix} \lambda_{\bar{y}} I_{n_a} & 0 & 0 \\ 0 & \lambda_{\bar{u}} I_{n_b} & 0 \\ 0 & 0 & \lambda_{\bar{y}} \end{pmatrix} \bar{\theta}$$

Puis, en injectant (17) et (18) dans cette dernière équation, et utilisant le fait que $a_{n_a+1} = 0$, on obtient :

$$(a_{n_a+1})_{LS}^* = \delta^{-1} \mathbf{v}^T R_{\phi\phi}^{-1} \begin{pmatrix} \lambda_{\bar{y}} I_{n_a} & 0 \\ 0 & \lambda_{\bar{u}} I_{n_b} \end{pmatrix} \theta \quad (20)$$

En combinant les équations (19) et (20) il vient :

$$\mathbf{v}^T R_{\phi\phi}^{-1} \begin{pmatrix} \lambda_{\bar{y}} a \\ \lambda_{\bar{u}} b \end{pmatrix} = r_{yy}(n_a+1) - \mathbf{v}^T R_{\phi\phi}^{-1} r_{\phi y}$$

$$\iff \lambda_{\bar{y}} \mathbf{v}^T M_{n_a} a + \lambda_{\bar{u}} \mathbf{v}^T M_{n_b} b = r_{yy}(n_a+1) - \mathbf{v}^T \theta_{LS}^* \quad (21)$$

où $M_{n_a} \in \mathbb{R}^{(n_a+n_b) \times n_a}$ et $M_{n_b} \in \mathbb{R}^{(n_a+n_b) \times n_b}$ dénotent respectivement les matrices formées en ne prenant que les n_a premières colonnes et les n_b dernières colonnes de la matrice $R_{\varphi\varphi}^{-1} \in \mathbb{R}^{(n_a+n_b) \times (n_a+n_b)}$.

L'algorithme BELS1-II est alors obtenu en utilisant les équations (15) et (21) pour estimer les variances des bruits, et se résume ainsi [10] :

1. estimation du vecteur des paramètres θ à partir des données disponibles $\{u(t), y(t) | t = 1, \dots, N\}$ par la méthode des moindres carrés, ce qui donne le vecteur des paramètres initial $\hat{\theta}_{BELS}^{(0)}(N) = \hat{\theta}_{LS}(N)$;
2. calcul de $\hat{\lambda}_{\bar{y}}^{(i+1)}$ et $\hat{\lambda}_{\bar{u}}^{(i+1)}$ en résolvant le système suivant, où les quantités asymptotiques ont été remplacées par leur estimation obtenue pour le nombre fini de données disponibles :

$$\begin{pmatrix} \hat{f}(\hat{\theta}_{LS}(N)) \\ \hat{r}_{yy}(n_a+1) - \mathbf{v}^T \hat{\theta}_{LS}(N) \end{pmatrix} = \Delta \begin{pmatrix} \hat{\lambda}_{\bar{y}}^{(i+1)} \\ \hat{\lambda}_{\bar{u}}^{(i+1)} \end{pmatrix}$$

avec :

$$\Delta \triangleq \begin{pmatrix} 1 + \hat{a}_{LS}^T(N) \hat{a}_{BELS}^{(i)}(N) & \hat{b}_{LS}^T(N) \hat{b}_{BELS}^{(i)}(N) \\ \mathbf{v}^T \hat{M}_{n_a} \hat{a}_{BELS}^{(i)}(N) & \mathbf{v}^T \hat{M}_{n_b} \hat{b}_{BELS}^{(i)}(N) \end{pmatrix}$$

3. calcul de $\hat{\theta}_{BELS}^{(i+1)}(N)$:

$$\begin{aligned} \hat{\theta}_{BELS}^{(i+1)}(N) &= \hat{\theta}_{LS}(N) \\ &+ \hat{R}_{\varphi\varphi}^{-1}(N) \begin{pmatrix} \hat{\lambda}_{\bar{y}}^{(i+1)} I_{n_a} & 0 \\ 0 & \hat{\lambda}_{\bar{u}}^{(i+1)} I_{n_b} \end{pmatrix} \hat{\theta}_{BELS}^{(i)}(N) \end{aligned}$$

4. répéter les étapes 2 et 3 jusqu'à convergence (selon le critère d'arrêt choisi).

Il va sans dire que les performances de ces algorithmes qui reposent sur la procédure itérative d'élimination de biais (7) dépendent grandement de la qualité des estimées de $\lambda_{\bar{u}}$ et $\lambda_{\bar{y}}$. Jusqu'à présent l'estimation était réalisée en résolvant un système linéaire de deux équations faisant intervenir ces deux variances.

Cependant, W.X. Zheng a observé que dans un certain nombre de cas les algorithmes BELS1 et BELS1-II requièrent un nombre plus important d'itérations pour atteindre le critère d'arrêt [11]. Or un critère d'évaluation d'un algorithme itératif est sa vitesse de convergence : il est toujours préférable qu'elle soit grande pour qu'il soit applicable en pratique.

Une autre constatation est que la procédure d'estimation des méthodes BELS1 et BELS1-II ne donne pas d'estimations satisfaisantes de $\lambda_{\bar{u}}$ et $\lambda_{\bar{y}}$ quand le rapport signal/bruit (RSB) est trop faible. La qualité de l'estimation des paramètres en est alors affectée dans le sens où la variance sur ces estimations est plus grande, bien qu'elles soient toujours non biaisées.

La méthode FBELS (pour Fast BELS), qui va être exposée dans la suite, repose sur l'idée suivante : utiliser non plus une seule mesure supplémentaire de la sortie, mais plusieurs, de manière à avoir plus de deux équations linéaires en $\lambda_{\bar{u}}$ et $\lambda_{\bar{y}}$. La résolution au sens des moindres carrés du système linéaire

sur-déterminé ainsi obtenu donnera, en moins d'itérations, une estimation satisfaisante des deux variances. De plus, les estimées des variances étant meilleures, les variances sur les paramètres se verront diminuées, et le nouvel algorithme sera plus fiable dans le cas de faibles RSB [11].

MÉTHODE FBELS

On augmente le degré du dénominateur de $G(q, \theta)$ non plus de un, mais d'un nombre $d > 1$, les nouveaux paramètres ainsi introduits ayant une valeur vraie égale à zéro. On définit l'opérateur de transfert augmenté et le vecteur des paramètres augmenté par :

$$\begin{aligned} \bar{G}(q, \theta) &= \frac{B(q^{-1}, \theta)}{\bar{A}(q^{-1}, \theta)} = \frac{b_1 q^{-1} + \dots + b_{n_b} q^{-n_b}}{1 + a_1 q^{-1} + \dots + a_{n_a+d} q^{-n_a-d}} \\ \bar{\theta}^T &= [\theta^T \ \alpha^T] ; \quad \alpha^T = [a_{n_a+1} \ \dots \ a_{n_a+d}] \end{aligned}$$

Pour le système augmenté, le vecteur de régression s'écrit :

$$\begin{aligned} \bar{\varphi}^T(t) &= [\varphi^T(t) \ \mu^T(t)] \\ \mu^T(t) &= [-y(t-n_a-1) \ \dots \ -y(t-n_a-d)] \end{aligned}$$

A nouveau, on constate que l'augmentation du degré du dénominateur de $G(q, \theta)$ de d revient à utiliser d mesures supplémentaires de la sortie.

L'estimée au sens des moindres carrés s'écrit :

$$\bar{\theta}_{LS}^* = R_{\bar{\varphi}\bar{\varphi}}^{-1} r_{\bar{\varphi}y} = \bar{\theta} - \bar{R}_{\bar{\varphi}\bar{\varphi}}^{-1} \begin{pmatrix} \lambda_{\bar{y}} I_{n_a} & 0 & 0 \\ 0 & \lambda_{\bar{u}} I_{n_b} & 0 \\ 0 & 0 & \lambda_{\bar{y}} I_d \end{pmatrix} \bar{\theta}$$

De même qu'auparavant, où une équation linéaire en $\lambda_{\bar{u}}$ et $\lambda_{\bar{y}}$ était déduite de l'ajout d'un nouveau paramètre, il est ici possible de déduire d équations supplémentaires, comme il est précisé dans le théorème suivant (dont la preuve se trouve en annexe A) :

THÉORÈME 1

Les variances des bruits $\lambda_{\bar{u}}$ et $\lambda_{\bar{y}}$ satisfont

$$\lambda_{\bar{y}} R_{\varphi\mu}^T M_{n_a} a + \lambda_{\bar{u}} R_{\varphi\mu}^T M_{n_b} b = r_{\mu y} - R_{\varphi\mu}^T \theta_{LS}^* \quad (22)$$

où $R_{\varphi\mu} = \bar{E} [\varphi(t) \mu(t)] \in \mathbb{R}^{(n_a+n_b) \times d}$, $r_{\mu y} = \bar{E} [\mu(t) y(t)]$, et $M_{n_a} \in \mathbb{R}^{(n_a+n_b) \times n_a}$ et $M_{n_b} \in \mathbb{R}^{(n_a+n_b) \times n_b}$ dénotent respectivement les matrices formées en ne prenant que les n_a premières colonnes et les n_b dernières colonnes de la matrice $R_{\varphi\varphi}^{-1} \in \mathbb{R}^{(n_a+n_b) \times (n_a+n_b)}$.

En utilisant (15) et (22) on obtient désormais un système de $(d+1)$ équations linéaires. La résolution au sens des moindres carrés de ce système sur-déterminé donne alors une estimation des deux variances $\lambda_{\bar{u}}$ et $\lambda_{\bar{y}}$. L'algorithme FBELS peut être vu comme une extension de la méthode BELS1-II, dans le sens où l'équation (22) pour $d=1$ devient la même que l'équation (21). Toutefois, dans ce nouvel algorithme, d mesures supplémentaires de la sortie sont utilisées et donc, lorsque $d > 1$, les deux variances sont estimées avec plus d'informations qu'auparavant.

Le nombre d est à choisir par l'utilisateur, mais le meilleur compromis entre temps de calcul et qualité des estimées semble obtenu pour $d=2$ ou 3 ; d'après différents tests pratiques, de plus grandes valeurs n'offrent pas de différence significative en terme de précision des estimées, tandis que le temps de calcul est accru [11].

Méthode	N	$a_1=-1.5$	$a_2=0.7$	$b_1=1.0$	$b_2=0.5$	$\lambda_{\bar{u}}=0.26$	$\lambda_{\bar{y}}=1.9$	EQMR	#Itér.
IV4	1000	-1.4757 ± 0.0347	0.6609 ± 0.0290	0.8015 ± 0.0381	0.3662 ± 0.0609	–	–	12.92%	–
	5000	-1.4960 ± 0.0126	0.6791 ± 0.0106	0.8124 ± 0.0170	0.3544 ± 0.0247	–	–	12.06%	–
OE	1000	-1.4980 ± 0.0345	0.6794 ± 0.0278	0.8528 ± 0.0462	0.2597 ± 0.0739	–	–	14.97%	3.526
	5000	-1.5041 ± 0.0135	0.6874 ± 0.0110	0.8805 ± 0.0213	0.2689 ± 0.0306	–	–	13.20%	2.786
BELS1	1000	-1.5009 ± 0.0598	0.7015 ± 0.0468	1.0046 ± 0.1132	0.5044 ± 0.1207	0.2458 ± 0.1429	1.8881 ± 0.1204	9.11%	6.216
	5000	-1.5000 ± 0.0253	0.7001 ± 0.0193	1.0006 ± 0.0497	0.5020 ± 0.0524	0.2575 ± 0.0605	1.8987 ± 0.0537	3.95%	6.880
BELS1-II	1000	-1.5010 ± 0.0598	0.7016 ± 0.0469	1.0050 ± 0.1131	0.5047 ± 0.1205	0.2462 ± 0.1423	1.8878 ± 0.1202	9.10%	6.060
	5000	-1.5000 ± 0.0253	0.7001 ± 0.0193	1.0005 ± 0.0498	0.5019 ± 0.0524	0.2573 ± 0.0606	1.8989 ± 0.0538	3.95%	6.000
FBELS ($d=3$)	1000	-1.4988 ± 0.0344	0.6996 ± 0.0269	1.0061 ± 0.0936	0.5054 ± 0.0806	0.2573 ± 0.0961	1.8871 ± 0.1064	6.57%	5.972
	5000	-1.5001 ± 0.0143	0.7002 ± 0.0108	1.0006 ± 0.0401	0.5017 ± 0.0343	0.2592 ± 0.0422	1.8992 ± 0.0474	2.79%	5.956

TAB. 1 – Résultats des simulations de Monte Carlo (500 réalisations, RSBE = RSBS \simeq 10 dB).

SIMULATIONS NUMÉRIQUES

L'objectif de cette section est de comparer les performances des trois algorithmes étudiés à l'aide de simulations de Monte Carlo de $M = 500$ réalisations. Le système considéré, ainsi que les conditions de simulation s'inspirent de [10] :

$$\mathcal{G}(\theta) : G_0(q) = \frac{1.0q^{-1} + 0.5q^{-2}}{1 - 1.5q^{-1} + 0.7q^{-2}}$$

Le signal d'entrée non bruité est généré comme la moyenne ajustée d'un bruit blanc $e(t)$ de variance unité :

$$u_0(t) = (1 - 0.3q^{-1} + 0.5q^{-2} - 0.7q^{-3} + 0.9q^{-4})e(t)$$

Les variances des bruits blancs en entrée et en sortie ont été fixées à $\lambda_{\bar{u}} = 0.26$ et $\lambda_{\bar{y}} = 1.9$ pour obtenir les rapports signal/bruit en entrée et en sortie suivants :

$$\text{RSBE} \triangleq 10 \log_{10} (P_{u_0} / P_{\bar{u}}) \simeq 10 \text{ dB}$$

$$\text{RSBS} \triangleq 10 \log_{10} (P_{y_0} / P_{\bar{y}}) \simeq 10 \text{ dB}$$

où P_x représente la puissance moyenne du signal x . Deux simulations de Monte Carlo ont été réalisées, l'une avec $N = 1000$ données d'entrée/sortie, l'autre avec $N = 5000$. Le critère d'arrêt adopté pour déterminer la convergence de chacun des algorithmes est :

$$\| \hat{\theta}_{BELS}^{(i+1)}(N) - \hat{\theta}_{BELS}^{(i)}(N) \| / \| \hat{\theta}_{BELS}^{(i)}(N) \| < 10^{-3}$$

Les résultats des simulations de Monte Carlo sont confinés dans le tableau 1, au sein duquel sont présentés : les valeurs moyennes des paramètres estimés, leur écart-type, les estimées des variances $\lambda_{\bar{u}}$ et $\lambda_{\bar{y}}$, le nombre moyen d'itérations avant convergence, ainsi que l'erreur quadratique moyenne relative sur les paramètres, définie par :

$$\text{EQMR} \triangleq \sqrt{\frac{1}{M} \sum_{j=1}^M \frac{\| \hat{\theta}_{BELS}^j(N) - \theta \|^2}{\| \theta \|^2}}$$

où $\hat{\theta}_{BELS}^j(N)$ représente la j -ième estimée de la simulation de Monte Carlo. Les système a également été identifié en utilisant les approches classiques oe et iv4, dédiées à l'identification de systèmes ayant uniquement un bruit en sortie ; le but étant de montrer l'intérêt des approches de compensation de biais analysées dans cet article.

La théorie, qui affirme que les trois algorithmes étudiés fournissent des estimées asymptotiquement non biaisées, se voit

confirmée par les simulations de Monte Carlo. En effet, ils affichent des performances peu ou prou similaires en terme de valeur moyenne des paramètres estimés : toutes les estimées sont très proches des valeurs vraies, les légers écarts n'étant que la traduction du fait que l'on ne dispose que d'un nombre *fini* de données. Les critères appropriés pour jauger des performances respectives des trois algorithmes sont donc les écarts-types, le nombre moyen d'itérations avant convergence, et le coût de calcul par itération.

Tout d'abord, regardons plus attentivement les algorithmes BELS1 et BELS1-II : leurs résultats sont identiques tant au niveau de l'écart-type qu'au niveau de la moyenne des estimées. En revanche, comme on l'a vu dans l'étude ci-avant, l'algorithme BELS1 fait intervenir le système augmenté, tandis que l'algorithme BELS1-II ne l'utilise que pour déduire les équations nécessaires à l'estimation des variances des bruits. Le coût algorithmique s'en trouve par conséquent réduit. Une autre distinction entre les deux vient du nombre moyen d'itérations : on voit que l'algorithme BELS1-II converge plus rapidement ; par conséquent, le coût algorithmique s'en trouve à nouveau réduit. L'algorithme BELS1-II est par conséquent strictement meilleur.

L'algorithme FBELS affiche quant à lui de meilleures performances : si le nombre moyen d'itérations ne se voit qu'un peu diminué, les écarts-types sur les paramètres sont en revanche significativement réduits. On voit donc que l'utilisation de plus de mesures de la sortie offre logiquement une amélioration de l'estimation.

La clé de voûte de ces algorithmes est l'estimation des variances $\lambda_{\bar{u}}$ et $\lambda_{\bar{y}}$: dans la procédure itérative d'élimination de biais (7), commune aux trois algorithmes, on remarque qu'il n'y a que cette estimation qui permette une amélioration. La qualité de l'estimée de la matrice de covariance $\hat{R}_{\phi\phi}(N)$ n'est en effet que fonction du nombre N de données disponibles. Donc : meilleures sont les estimées des variances $\lambda_{\bar{u}}$ et $\lambda_{\bar{y}}$, meilleures seront les estimées des paramètres, comme en attestent les résultats contenus dans le tableau 1.

Les méthodes oe et iv4 n'offrent pas d'estimation convergente dans un contexte erreurs-en-les-variables. L'intérêt de développer des méthodes spécifiques pour ce problème est mis en exergue : si les estimées des paramètres de $A_0(q^{-1})$ sont satisfaisantes, celles des paramètres de $B_0(q^{-1})$ sont en revanche fortement biaisées. Cette observation se voit confirmée par les résultats du tableau 2 : si l'on diminue le RSB en entrée jusqu'à 5 dB, les résultats de oe et iv4 ne sont plus acceptables, tandis que FBELS donne encore de bonnes estimées.

Méthode	oe	iv4	FBELS ($d=3$)
$a_1=-1.5$	-1.5124 ± 0.0551	-1.4574 ± 0.0588	-1.4973 ± 0.0459
$a_2=0.7$	0.6780 ± 0.0461	0.6295 ± 0.0516	0.6983 ± 0.0397
$b_1=1.0$	0.6678 ± 0.0495	0.5856 ± 0.0362	1.0083 ± 0.0997
$b_2=0.5$	0.0747 ± 0.0872	0.2306 ± 0.0642	0.5044 ± 0.0898
EQMR	27.74%	25.65%	9.46%

TAB. 2 – Résultats des simulations de Monte Carlo (500 réalisations, $N = 1000$, RSBE $\simeq 5$ dB, RSBS $\simeq 10$ dB).

CONCLUSION ET PERSPECTIVES

Trois méthodes de compensation de biais pour l'identification de systèmes linéaires invariants avec erreurs-en-les-variables ont été analysées. Elles sont fondées sur l'utilisation d'une ou plusieurs mesures supplémentaires de la sortie pour obtenir des équations linéaires permettant d'estimer les variances des bruits. Lorsque les bruits en entrée et en sortie sont blancs, celles-ci déterminent entièrement le biais de la méthode des moindres carrés, et leurs estimées peuvent donc être utilisées pour obtenir des estimées non biaisées des paramètres. Des simulations numériques ont permis d'évaluer et comparer les performances des trois algorithmes, qui se sont révélées satisfaisantes.

Cet article constitue un travail préliminaire qui permet de dégager les perspectives les suivantes :

- les méthodes considérées dans cet article ne sont applicables que dans le cas où les bruits en entrée et en sortie sont blancs, hypothèse restrictive qui implique notamment que le système n'opère pas en boucle fermée. Il serait donc intéressant d'étendre ces méthodes au cas où les bruits sont colorés. Les travaux récents de W.X. Zheng [12] considèrent le cas d'un bruit blanc en entrée et coloré en sortie, mais le cas de bruits colorés en entrée *et* en sortie reste encore un problème ouvert ;
- comparer les méthodes d'élimination de biais à d'autres méthodes existantes ;
- l'extension des approches de compensation de biais développées dans cet article au cas de l'identification de modèles à temps continu.

Par ailleurs, la méthode de Koopmans-Levin [1] a pour désavantage de requérir la connaissance *a priori* du rapport des variances des bruits. Il serait intéressant d'utiliser les variances estimées par la méthode FBELS pour s'affranchir de cet inconvénient, et comparer les performances des méthodes FBELS et Koopmans-Levin.

A. PREUVE DU THÉORÈME

La matrice de covariance du vecteur de régression augmenté $\bar{\varphi}(t)$ peut être décomposée en :

$$R_{\bar{\varphi}\bar{\varphi}} = \begin{bmatrix} R_{\varphi\varphi} & R_{\varphi\mu} \\ R_{\varphi\mu}^T & R_{\mu\mu} \end{bmatrix}$$

L'inversion de cette matrice nous donne :

$$R_{\bar{\varphi}\bar{\varphi}}^{-1} = \begin{bmatrix} R_{\varphi\varphi}^{-1} + R_{\varphi\varphi}^{-1} R_{\varphi\mu} M R_{\mu\mu}^T R_{\varphi\varphi}^{-1} & -R_{\varphi\varphi}^{-1} R_{\varphi\mu} M^{-1} \\ -M^{-1} R_{\varphi\mu}^T R_{\varphi\varphi}^{-1} & M^{-1} \end{bmatrix}$$

où $M \triangleq R_{\mu\mu} - R_{\varphi\mu}^T R_{\varphi\varphi}^{-1} R_{\varphi\mu}$

Puis on remplace $R_{\bar{\varphi}\bar{\varphi}}^{-1}$ par son expression ainsi obtenue dans les deux équations suivantes :

$$\begin{cases} \bar{\theta}_{LS}^* = \begin{pmatrix} \theta_{LS}^* \\ \alpha_{LS}^* \end{pmatrix} = \bar{\theta} - \bar{R}_{\bar{\varphi}\bar{\varphi}}^{-1} \begin{pmatrix} \lambda_{\bar{y}} I_{n_a} & 0 & 0 \\ 0 & \lambda_{\bar{u}} I_{n_b} & 0 \\ 0 & 0 & \lambda_{\bar{y}} I_d \end{pmatrix} \bar{\theta} \\ \bar{\theta}_{LS}^* = R_{\bar{\varphi}\bar{\varphi}}^{-1} \bar{E} [\bar{\varphi}(t)y(t)] \end{cases}$$

On obtient respectivement :

$$\begin{cases} \alpha_{LS}^* = M^{-1} R_{\varphi\mu}^T R_{\varphi\varphi}^{-1} \begin{pmatrix} \lambda_{\bar{y}} I_{n_a} & 0 \\ 0 & \lambda_{\bar{u}} I_{n_b} \end{pmatrix} \theta \\ \alpha_{LS}^* = -M^{-1} R_{\varphi\mu}^T R_{\varphi\varphi}^{-1} \bar{E} [\varphi(t)y(t)] + M^{-1} \bar{E} [\mu(t)y(t)] \end{cases}$$

Ces deux équations permettent alors d'écrire :

$$\begin{aligned} R_{\varphi\mu}^T R_{\varphi\varphi}^{-1} \begin{pmatrix} \lambda_{\bar{y}} a \\ \lambda_{\bar{u}} b \end{pmatrix} &= \bar{E} [\mu(t)y(t)] - R_{\varphi\mu}^T R_{\varphi\varphi}^{-1} \bar{E} [\varphi(t)y(t)] \\ \iff \lambda_{\bar{y}} R_{\varphi\mu}^T M_{n_a} a + \lambda_{\bar{u}} R_{\varphi\mu}^T M_{n_b} b &= \bar{E} [\mu(t)y(t)] - R_{\varphi\mu}^T \theta_{LS}^* \end{aligned}$$

ce qui complète la preuve.

REFERENCES

- [1] K.V. Fernando and H. Nicholson. Identification of linear systems with input and output noise : the Koopmans-Levin method. *IEE Proceedings, Part D*, 132(1) :30–36, 1985.
- [2] H. Garnier, M. Gilson, and W.X. Zheng. A bias-eliminated least-squares method for continuous-time model identification of closed-loop systems. *International Journal of Control*, 73(1) :38–48, 2000.
- [3] M. Gilson and P. Van den Hof. On the relation between a bias-eliminated least-squares (BELS) and an IV estimator in closed-loop identification. *Automatica*, 37(10) :1593–1600, 2001.
- [4] I.D. Landau and A. Besançon-Voda. *Identification des systèmes*. Traité IC2 – section systèmes automatisés. Hermès, Paris, 2001.
- [5] L. Ljung. *System Identification : Theory for the User*. Prentice Hall, 2nd edition, 1999.
- [6] T. Söderström. Why are errors-in-variables problems often tricky ? *Technical report 2002-040*, 2002.
- [7] S. Van Huffel. *Total least squares and errors-in-variables modeling*. Kluwer Academic Publishers, 2002.
- [8] E. Walter and L. Pronzato. *Identification de modèles paramétriques à partir de données expérimentales*. Masson, 1994.
- [9] W.X. Zheng. Transfer function estimation from noisy input and output data. *International Journal of Adaptive Control and Signal Processing*, 12(4) :365–380, 1998.
- [10] W.X. Zheng. On least-squares identification of stochastic linear systems with noisy input-output data. *International Journal of Adaptive Control and Signal Processing*, 13(3) :131–143, 1999.
- [11] W.X. Zheng. Unbiased identification of stochastic linear systems from noisy input and output measurements. In *39th IEEE Conference on Decision and Control*, volume 3, pages 2710–2715, Sydney, Australia, December 2000.
- [12] W.X. Zheng. A bias correction method for identification of linear dynamic errors-in-variables models. *IEEE Transactions on Automatic Control*, 47(7) :1142–1147, July 2002.