Modèles réduits ARX et produit de convolution en thermique linéaire des systèmes invariants.

Reduced ARX models and convolution product for heat transfer in times invariant systems.

Denis MAILLET[1*], Célien ZACHARIE[1], Benjamin RÉMY[1]

[1]Laboratoire Énergies et Mécanique Théorique et Appliquée
Université de Lorraine, CNRS, 2 avenue de la Forêt de Haye - BP 90161 - Vandoeuvre-lès-Nancy, France.
[*](auteur correspondant : denis.maillet@univ-lorraine.fr)

Résumé – Les modèles paramétriques polynomiaux de structure ARX sont de plus en plus employés en caractérisation des transferts de chaleur en thermique des systèmes linéaires et invariants, du fait de leur robustesse dans les problèmes d'inversion, c'est-à-dire dans les problèmes d'identification de modèle ou d'estimation d'entrée. Leur parcimonie permet d'obtenir en effet des résidus très faibles avec un nombre restreint de coefficients. Cet article montre, sur une base théorique algébrique, que ces modèles ARX peuvent être déduits des modèles convolutifs.

Mots-clés : Calibration ; Identification ; Problème inverse ; Produit de convolution ; Parcimonie.

Abstract – Parametric polynomial models of ARX structure are becoming very popular for characterizing heat transfer for linear time independent systems, because of their robustness in inversion problems such as system identification or input estimation. Their parsimonious character allows low level residuals with a small number of parameters. This paper shows, on a theoretical algebraic basis, that these ARX models can be derived from the convective models.

Keywords: Calibration; Identification; Problème inverse; Convolution product; Parsimony.

1. Introduction

L'objet de ce papier est de faire le lien entre les modèles réduits de structure ARX et les modèles convolutifs qui relient une réponse en température $y\,(t)$ en un point donné de l'espace à une source $u\,(t)$, ici une excitation thermique transitoire $u\,(t)$, qui peut être soit une autre température, ou une puissance thermique dissipée, locale ponctuelle ou répartie. Ce lien causal entre une entrée et une sortie est caractérisé, au travers d'un produit de convolution, par une réponse impulsionnelle $h\,(t)$. Ce modèle convolutif est toujours applicable en dynamique des systèmes Linaires à coefficients Invariants en Temps (LIT) pourvu que l'excitation soit unique et séparable entre temps et espace [1, 2]. Il s'écrit :

$$y(t) = (h * u)\,(t) = \int_0^t h\,(t')\,u\,(t-t')\,\mathrm{d}t' \quad \Leftrightarrow \quad \bar{y}(p) = \bar{h}\,(p)\;\bar{u}\,(p) \qquad \text{(1a, b)}$$

Ici le symbole étoile désigne le produit de convolution, p est la variable de Laplace et la barre supérieure sur une fonction désigne sa transformée.

https://doi.org/10.25855/SFT2021-049

Un des problèmes inverses liés à ce modèle est celui de l'identification de $h\,(t)$ à partir d'observations ou de mesures de $u\,(t)$ et de $y\,(t)$ en des instants discrets de l'intervalle $\left]\,0\,,t_f\,\right]$.

A cause de la symétrie entre h et u dans (1a) ou (1b), ce problème spécifique de thermique inverse est différent du problème classique de conduction inverse (IHCP) où $u\,(t)$ est recherché à partir de valeurs échantillonnées de $y\,(t)$ qui sont entachées d'un bruit de mesure et dont la réponse impulsionnelle exacte $h\,(t)$ est connue, pour les raisons suivantes :

i) dans ce problème d'identification non seulement $y\,(t)$ mais aussi $u\,(t)$ est entachée d'un bruit,
ii) la réponse impulsionelle (une transmittance si u est une température, ou une impédance s'il s'agit d'une puissance thermique) a un support fini qui est inconnu.

Le point i) peut être résolu par une régularisation utilisant les moindres carrés totaux (TLS) [1]. Le point ii) est plus délicat: le choix de l'ensemble des temps de mesure, et donc du pas de temps, a une influence importante sur la résolution de la fonction à estimer (nombre et localisation temporelle de la base de fonctions servant à la paramétrisation de $h\,(t)$), même en l'absence de bruit. Donc, si l'accord entre le choix des instants de mesure et le support (inconnu) de $h\,(t)$ n'est pas bon, une régularisation classique (Tikhonov) ou par TSVD (décomposition en valeurs singulières tronquée) de l'opérateur du modèle convolutif (1a) mis sous une forme discrète matricielle utilisant une matrice carrée (autant d'inconnues que de temps de mesure) ne peut pas toujours fournir une estimation satisfaisante, c'est-à-dire qui respecte le principe de non contradiction (« discrepancy principle ») de Morozov sur les résidus.

Une solution alternative de ce problème IHCP spécifique consiste à utiliser un modèle paramétrique polynomial spécifique, de structure AutoRégressive à entrée eXterne (ARX) dont la structure est donnée par l'équation (2) :

$$y_k = -\sum_{i=1}^{na} a_i\, y_{k-i} + \sum_{j=1}^{nb} b_j\, u_{k-j-nk} \tag{2}$$

La sortie du modèle $y_k = y\,(t_k)$ à un instant $t_k = k\,\Delta t$ courant est ici une combinaison des n_a sorties précédentes et des n_b entrées précédentes, avec un éventuel retard temporel égal à $n_k\,\Delta t$, où Δt est le pas d'acquisition ou d'observation [3]. Une identification par moindres carrés sur tout le domaine temporel de ce modèle discret, c'est-à-dire une estimation des $(n_a + n_b)$ paramètres a_i et b_j, pour un choix optimum du triplet $(n_a\,, n_b, n_k)$ qui minimise les résidus correspondants, est généralement assez robuste [4]. Ce type d'algorithme est par exemple disponible dans l'environnement Matlab. Des applications récentes [5, 6] dans le domaine de la thermique des systèmes, dont la modélisation est difficile, commencent à apparaitre.

Dans cet article, les modèles convolutif (1a) et ARX sont comparés sur une base théorique, en utilisant une paramétrisation de chacune des trois fonctions de (1a) sur un peigne de Dirac. Ceci conduit à des expressions explicites des paramètres du modèle ARX en fonction de la réponse impulsionnelle $h\,(t)$.

2. Interprétation des coefficients ARX en fonction de la réponse impulsionnelle

2.1 Version discrète du modèle convolutif

En pratique, comme vu plus haut, les observations des valeurs instantanées du signal de sortie $y\,(t)$ sont faites pour des temps discrets isochrones et sont donc échantillonnées avec un pas temporel à partir de cette fonction de sortie à temps continu. Ces valeurs sont calculées grâce à une quadrature numérique (1a) de l'intégrale de convolution :

$$y(t_k) = \int_0^{t_k} h\,(t')\,u\,(t_k - t')\,\mathrm{d}t' = \sum_{j=1}^{k} \int_{t_{j-1}}^{t_j} h\,(t')\,u\,(t_k - t')\,\mathrm{d}t' \approx \sum_{j=1}^{k} \Delta t\ \tilde{h}_j\ \tilde{u}_{k-j+1} \qquad \text{(3a)}$$

Ici la notation tilde surmontant les symboles de la réponse impulsionnelle et de l'entrée désigne les moyennes glissantes de ces fonctions sur une fenêtre temporelle de largeur Δt. On les évalue aux bornes temporelles supérieures de ces intervalles désignés par leur indice :

$$\tilde{h}_j = \tilde{h}\,(t_j) \equiv \frac{1}{\Delta t}\int_{t_{j-1}}^{t_j} h\,(t')\ \mathrm{d}t' \quad \text{avec} \quad \tilde{x}\,(t) \equiv \frac{1}{\Delta t}\int_{t-\Delta t}^{t} x\,(t')\ \mathrm{d}t' \quad \text{pour} \quad x = h \ \text{ or } \ u$$

$$\text{et} \quad \tilde{u}_{k-j+1} = \tilde{u}\,(t_{k-j+1}) = \frac{1}{\Delta t}\int_{t_{k-j}}^{t_{k-j+1}} u\,(t'')\ \mathrm{d}t'' = \frac{1}{\Delta t}\int_{t_{j-1}}^{t_j} u\,(t_k - t')\ \mathrm{d}t' \qquad \text{(3b)}$$

L'équation (3a), qui est une approximation dans le cas général, devient exacte si la covariance déterministe de $h\,(t)$ et $u_k^{flip}(t) = u\,(t_k - t)$ est égale à zéro sur chacun de leurs intervalles communs $\Delta_j = \left]\, t_{j-1}\,,\, t_j \right]$. C'est par exemple le cas si les deux fonctions sont strictement égales à leurs versions paramétrisées sur une base de fonctions constantes par morceaux, ou si elles varient linéairement sur chaque intervalle temporel.

Il est très intéressant d'introduire ici les notions de doses de ces fonctions sur chaque intervalle Δ_j :

$$\hat{x}_j = \int_{t_{j-1}}^{t_j} x\,(t')\ \mathrm{d}t' = \tilde{x}_j\,\Delta t \quad \text{pour} \quad x = h \ \text{ ou } \ u \ \Rightarrow \ y_k = y(t_k) = \sum_{j=1}^{k} \hat{h}_j\ \tilde{u}_{k-j+1} = \sum_{j=1}^{k} \tilde{h}_j\ \hat{u}_{k-j+1} \qquad \text{(4a, b)}$$

Une hypothèse forte est ici posée: on suppose que la fonction $h\,(t)$ est nulle pour des instants postérieurs à un temps $t_{nh} = n_h\,\Delta t$, ce qui suppose que le support de cette fonction est inclus dans l'intervalle $\left[\ 0\,,\, t_{nh}\ \right]$ avec également $h\,(0) = 0$. Donc le nombre minimum m d'observations de $u\,(t)$ et $y\,(t)$ nécessaire à une reconstruction de $h\,(t)$ est égal à n_h. Nous supposons donc $m \geq n_h$.

Une représentation matricielle du modèle discret (3a) existe, si on considère les différentes variables, c'est à dire les valeurs échantillonnées de $h\,(t)$ et de $u\,(t)$ à la fin de chacun des m intervalles Δ_j ainsi que les doses correspondantes. Cette représentation est donnée ici pour un instant t_k appartenant à l'intervalle des observations $\left[\, t_1\,,\, t_m \right]$:

$$\boldsymbol{y} = \mathbf{N}(\tilde{\boldsymbol{h}})\ \hat{\boldsymbol{u}} = \mathbf{N}(\tilde{\boldsymbol{u}})\ \hat{\boldsymbol{h}} \quad \text{with} \quad \hat{\boldsymbol{u}} \equiv \Delta t\ \tilde{\boldsymbol{u}} \quad \text{and} \quad \hat{\boldsymbol{h}} \equiv \Delta t\ \tilde{\boldsymbol{h}} \qquad \text{(5a, b)}$$

$$\text{où} \quad \boldsymbol{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_k \end{bmatrix} \quad ; \quad \mathbf{N}(\boldsymbol{x}) \equiv \begin{bmatrix} x_1 & & & & \\ x_2 & x_1 & & 0 & \\ x_3 & x_2 & x_1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ x_k & x_{k-1} & x_{k-2} & \cdots & x_1 \end{bmatrix} \quad \text{avec} \quad \boldsymbol{x} \equiv \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_k \end{bmatrix} \quad \text{pour} \quad x = \tilde{h} \text{ or } \tilde{u} \tag{5c}$$

Ici $\mathbf{N}(\boldsymbol{x})$ est une matrice de Toeplitz triangulaire inférieure dont les colonnes sont le vecteur $\boldsymbol{x}$, ainsi que ses versions décalées et tronquées.

Les deux formes du modèle convolutif discret indiquées plus haut peuvent être utilisées, dans le sens direct, pour simuler la sortie pour des observations ou des mesures connues de l'entrée si la réponse impulsionnelle est également connue.

Cependant, dans un problème de conduction inverse, ou même de conduction-advection inverse pour un système de type LIT, pour lequel dont on cherche à évaluer l'entrée, pour des valeurs discrètes connues de $y(t)$ et $h(t)$, éventuellement en présence d'un bruit additif, la forme (5a) est la plus adaptée, puisque le vecteur $\hat{\boldsymbol{u}}$ des doses de l'entrée est alors l'inconnue intrinsèque. Par contre, ceci nécessite d'accéder au vecteur $\hat{\boldsymbol{h}}$ des moyennes par intervalle de $h(t)$, ce qui ne peut être fait qu'à l'aide d'une moyenne arithmétique des valeurs de cette fonction aux bornes de chaque intervalle Δ_j, voir l'équation (6a).

D'une façon symétrique, dans le cas d'un problème inverse d'identification de $h(t)$ à partir de la connaissance des valeurs discrètes de $y(t)$ et de $h(t)$, la forme (5b) est plus appropriée. Dans ce cas, le vecteur $\hat{\boldsymbol{h}}$ doit être calculé en remplaçant le vecteur des moyennes $\tilde{\boldsymbol{u}}$ par un vecteur des moyennes arithmétiques correspondantes, voir l'équation (6b).

$$\tilde{\boldsymbol{x}} \approx \frac{1}{2}(\boldsymbol{x} + \boldsymbol{x}^{prior}) = \boldsymbol{K}\,\boldsymbol{x} \quad \text{with} \quad \boldsymbol{x}^{prior} \equiv \begin{bmatrix} x_0 = 0 \\ x_1 \\ x_2 \\ \vdots \\ x_{k-1} \end{bmatrix} \quad \text{for } x = h \text{ or } u \text{ and } \boldsymbol{K} \equiv \frac{1}{2}\begin{bmatrix} 1 & & & \\ 1 & 1 & & 0 \\ & \ddots & \ddots & \\ & 0 & \ddots & 1 \\ & & 1 & 1 \end{bmatrix} \tag{6a, b}$$

Pour les deux problèmes inverses en dimensions finies vus plus haut, les trois fonctions du modèle portant sur le produit de convolution continu (1a) sont remplacées par leurs formes paramétrisées qui peuvent être obtenues par projection i) soit sur une base de fonctions constantes par morceaux ou ii) sur une base de distributions de Dirac (impulsions). Les deux formes sont strictement équivalentes en ce qui concerne les valeurs discrètes de la sortie sur une grille temporelle, mais la deuxième, qui correspond à une projection sur un peigne de Dirac $C_{\Delta t}(t)$, permet d'obtenir des transformées de Laplace très simple des trois fonctions échantillonnées :

$$y_{param}(t) = \int_0^\infty C_{\Delta t}(t)\, y(t)\, \mathrm{d}t = \left(h_{param} * u_{param}\right)(t)$$
$$\text{où } C_{\Delta t}(t) = \sum_{i=1}^{\infty} \delta(t - i\,\Delta t) \quad \text{et} \quad x_{param}(t) = \int_0^\infty C_{\Delta t}(t)\, \tilde{x}(t_i)\, \mathrm{d}t \quad \text{pour } x = y,\ h \text{ ou } u \tag{7}$$

2.2 le modèle ARX et son interprétation à partir de sa version en temps continu

Les sorties y_k du modèle ARX (2) sont considérées ici comme des valeurs échantillonnées d'un modèle ARX à temps continu $y_{ARX}\ (t)$ défini par :

$$y_{ARX}\ (t) = -\sum_{i=1}^{na} a_i\ y_{ARX}\ (t-\tau_i) + \sum_{j=1}^{nb} b_j\ u\ (t-\tau_{j+nk}) \quad \text{où} \quad \tau_i = i\ \Delta t \tag{8}$$

Cette équation est écrite dans le domaine de Laplace, en utilisant la propriété des transformées des fonctions décalées en temps :

$$\overline{y}_{ARX}\ (p) = -\left(\sum_{i=1}^{na} a_i\ \exp(-\tau_i\ p)\right)\overline{y}_{ARX}\ (p) + \left(\sum_{j=1}^{nb} b_j\ \exp(-\tau_{j+nk}\ p)\right)\overline{u}\ (p) \tag{9}$$

Une comparaison des versions de Laplace du modèle convolutif(1b) et du modèle ARX à temps continu (9) donne :

$$\overline{h}\ (p) = \frac{\sum_{j=1+nk}^{nd} b_j\ \exp(-\tau_j\ p)}{1+\sum_{i=1}^{na} a_i\ \exp(-\tau_i\ p)} \quad \text{où} \quad nd = nk + nb \tag{10}$$

Les équations (1b) et (10), écrites pour une valeur nulle de la variable de Laplace p, fournissent des relations très intéressantes entre les doses cumulées des trois fonctions :

$$\overline{h}\ (0) = \frac{\overline{y}\ (0)}{\overline{u}\ (0)} = \frac{\sum_{i=1}^{m} \tilde{y}_i}{\sum_{i=1}^{m} \tilde{u}_i} = \frac{\sum_{j=1+nk}^{nd} b_j}{1+\sum_{i=1}^{na} a_i} \quad \Rightarrow \quad h^{ss} = \frac{y^{ss}}{u^{ss}} = \frac{\sum_{i=1}^{i_{sr}} y_i}{\sum_{i=1}^{i_{sr}} u_i} \tag{11a, b}$$

Dans l'équation (11b) u^{ss} désigne une entrée en régime permanent (steady state) et y^{ss} est la réponse permanente correspondante pour le système physique considéré caractérisé par sa réponse impulsionnelle $h\ (t)$. Nous supposons ici que $y\ (t)$ est une variation de température. Donc h^{ss} est une résistance thermique (en K/W) si $u\ (t)$ est une puissance thermique (en watts) et on peut l'appeler transmittance permanente (une grandeur adimensionnelle) si $u\ (t)$ est une variation de température qui génère $y\ (t)$ de manière causale.

En théorie, l'indice i_{sr} du temps final qui apparait dans la somme (11b) devrait être l'infini, puisqu'il correspond au temps nécessaire pour atteindre un régime permanent en entrée et en sortie, qui sont en fait des niveaux asymptotiques. Par contre, en pratique, le rapport $y\ (t)\,/\,h\ (t)$ peut atteindre le niveau h^{ss}, avec une précision acceptable, au bout d'un temps fini t_{sr} qui peut être défini comme le début d'un *régime glissant*.

2.3 Version matricielle du modèle ARX

Le modèle ARX (2) est écrit pour $na = nb = m-1$ et $nk = 0$, et un terme supplémentaire est introduit pour $j = 0$ dans la deuxième somme :

$$y_k = -\sum_{i=1}^{m-1} a_i\ y_{k-i} + \sum_{j=0}^{m-1} b_j\ u_{k-j} \quad \Rightarrow \quad \sum_{i=0}^{m-1} a_i\ y_{k-i} = \sum_{j=0}^{m-1} b_j\ u_{k-j} \quad \text{avec } a_0 = 1 \tag{12a, b}$$

La présence de ce terme en b_0 est expliquée plus loin. Un changement des indices i et j en $i-1$ et $j-1$ dans (12b) fournit l'équation (13a):

$$\sum_{i=1}^{m} a_{i-1}\; y_{k-i+1} = \sum_{j=1}^{m} b_{j-1}\; u_{k-j+1} \;\Rightarrow\; \sum_{i=1}^{k} a_{i-1}\; y_{k-i+1} = \sum_{j=1}^{k} b_{j-1}\; u_{k-j+1} \qquad \text{(13a, b)}$$

Le changement des bornes supérieures des sommes entre (13a) et (13b) provient des valeurs nulles de $u\,(t)$ and $y\,(t)$ pour des temps négatifs ou nuls. Cette équation, multipliée par le pas temporel Δt, est écrite pour les instants entre t_1 et t_k, et la forme du système d'équations qui en résulte correspond au produit de convolution entre deux vecteurs déjà utilisé dans les équations (5a à c) :

$$\mathbf{N}(\boldsymbol{a}^{prior})\; \boldsymbol{y} = \mathbf{N}(\boldsymbol{u})\; \boldsymbol{b}^{prior} \quad \text{avec } \boldsymbol{x}^{prior} = \begin{bmatrix} x_0 & x_1 & \cdots & x_{k-1} \end{bmatrix}^T \text{ pour } x = a \text{ ou } b \qquad \text{(13c)}$$

La notation $\boldsymbol{x}^T$ désigne ici la transpose d'un vecteur ou d'une matrice $\boldsymbol{x}$.

2.4 Introduction des coefficients a du modèle ARX dans le modèle convolutif

Le membre de gauche de l'équation (4b), qui est la réponse discrète $y_k = y(t_k)$ dans le produit de convolution, correspond à la réponse du modèle à une excitation forcée. Si on effectue une translation temporelle de l'échelle des temps dont l'origine passe de $t_0 = 0$ à t_r, la sortie y_k peut être considérée comme la somme de $y_r = y(t_r)$ et d'un terme forcé, le produit de convolution de la réponse impulsionnelle translatée $h\,(t-t_r)$ et de la partie restante de l'entrée originale, c'est-à-dire de $u\,(t)$ pour $t > t_r$:

$$y(t) = y\,(t_r) + y_r^f\,(t) \;\text{ avec }\; y_r^f\,(t) = \int_{t_r}^{t} h\,(t')\, u(t\text{-}t')\,\mathrm{d}t \quad \text{pour } r = 0 \text{ to } k-1 \qquad \text{(14a)}$$

La version discrète de cette décomposition s'écrit :

$$y_k = \Delta t \sum_{j=1}^{k} \tilde{h}_j\; \tilde{u}_{k-j+1} = y_r + y_r^f\,(t_k) \quad \text{avec} \quad y_r^f\,(t_k) \equiv \sum_{j=r+1}^{k} \tilde{h}_j\; \tilde{u}_{k-j+1} = \sum_{i=1}^{k-r} \tilde{h}_{k-i+1}\; \tilde{u}_i \qquad \text{(14b, c)}$$

$$\text{où } i = k - j + 1 \text{ pour } r = 0 \text{ à } k-1 \text{ et pour } k = 0 \text{ à } m-1$$

Chaque équation donnant y_k ci-dessus, correspond à un index r pour l'instant de transition t_r qui peut varier de 0 à $k-1$. Elle est multipliée par un coefficient α'_{k-r}, et les k équations résultantes sont ensuite additionnées pour ces différentes de r , ce qui donne :

$$\alpha'_{sum}\, y_k = \sum_{r=0}^{k-1} \alpha'_{k-r}\; y_r + \sum_{r=0}^{k-1} \alpha'_{k-r}\; y_r^f\,(t_k) \;\text{ avec }\; \alpha'_{sum} \equiv \sum_{r=0}^{m-1} \alpha'_{k-r} = \sum_{i=0}^{m-1} \alpha'_i \qquad \text{(15a)}$$

Une division des k équations ci-dessus par a'_{sum} fait apparaitre k nouveaux coefficients α_i :

$$y_k = -\sum_{r=0}^{k-1} \alpha_{k-r}\; y_r - \sum_{r=0}^{k-1} \alpha_{k-r}\; y_r^f\,(t_k) \quad \text{avec} \quad \alpha_i \equiv -\alpha'_i / \alpha'_{sum} \quad \text{pour } i = 1 \text{ à } k \qquad \text{(15b)}$$

Après introduction d'un coefficient α_0 et en utilisant le fait que $y_0 = 0$, l'équation(15b) devient :

$$\sum_{r=1}^{k} \alpha_{k-r}\; y_r = -\sum_{r=0}^{k-1} \alpha_{k-r}\; y_r^f\,(t_k) \qquad \text{avec} \quad \alpha_0 = 1 \qquad \text{(15c)}$$

Les changements d'indice $i = k - r + 1$ dans le membre de gauche et $j = k - r$ dans celui de droite, en notant de plus que $y_0^f(t_k)$ est égal à y_k, voir (14b), donnent :

$$\sum_{i=1}^{k} \alpha_{i-1}\ y_{k-i+1} = -\sum_{j=1}^{k} \alpha_j\ y_{k-j}^f\ (t_k) \tag{15d}$$

De la même manière que pour le passage de (13b) à (13c), l'équation (15d) est écrite sous forme matricielle :

$$\mathbf{N}(\boldsymbol{\alpha}^{prior})\ \boldsymbol{y} = -\mathbf{N}(\boldsymbol{\alpha})\ \boldsymbol{y}^f(t_k) \text{ avec } \boldsymbol{\alpha}^{prior} \equiv \begin{bmatrix} \alpha_0 = 1 & \alpha_1 & \cdots & \alpha_{k-1} \end{bmatrix}^T \text{ et } \boldsymbol{\alpha} \equiv \begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_k \end{bmatrix}^T \tag{16}$$

Le vecteur $\boldsymbol{y}^f(t_k)$, dont les composantes sont données par l'équation (14c), s'écrit :

$$\begin{bmatrix} y_0^f(t_k) \\ y_1^f(t_k) \\ \\ \\ y_{k-1}^f(t_k) \end{bmatrix} = \Delta t \begin{bmatrix} \tilde{u}_k & \tilde{u}_{k-1} & \cdots & \cdots & \tilde{u}_1 \\ & \tilde{u}_k & \tilde{u}_{k-1} & & \vdots \\ & & \ddots & \ddots & \vdots \\ & 0 & & \ddots & \tilde{u}_{k-1} \\ & & & & \tilde{u}_k \end{bmatrix} \begin{bmatrix} \tilde{h}_1 \\ \tilde{h}_2 \\ \vdots \\ \vdots \\ \tilde{h}_k \end{bmatrix} = \Delta t \begin{bmatrix} \tilde{h}_1 & \tilde{h}_2 & \cdots & \cdots & \tilde{h}_k \\ & \tilde{h}_1 & \tilde{h}_2 & & \vdots \\ & & \ddots & \ddots & \vdots \\ & 0 & & \ddots & \tilde{h}_2 \\ & & & & \tilde{h}_1 \end{bmatrix} \begin{bmatrix} \tilde{u}_k \\ \tilde{u}_{k-1} \\ \vdots \\ \vdots \\ \tilde{u}_{k-1} \end{bmatrix} \tag{17a}$$

ou, en utilisant l'opérateur matriciel de convolution ainsi qu'une matrice de permutation spécifique $\boldsymbol{J}$:

$$\boldsymbol{y}^f(t_k) = \mathbf{N}^T(\tilde{\boldsymbol{u}}^{flip})\ \hat{\boldsymbol{h}} = \mathbf{N}^T(\tilde{\boldsymbol{h}})\ \hat{\boldsymbol{u}}^{flip} \text{ avec } \boldsymbol{x}^{flip} \equiv \begin{bmatrix} x_k & x_{k-1} & \cdots & x_1 \end{bmatrix}^T = \boldsymbol{J}\,\boldsymbol{x} \text{ et } \boldsymbol{J} \equiv \begin{bmatrix} & & & 1 \\ & 0 & \iddots & \\ & 1 & 0 & \\ 1 & & & \end{bmatrix} \tag{17b}$$

La substitution de (17b) dans (16) donne :

$$\mathbf{N}(\boldsymbol{\alpha}^{prior})\ \boldsymbol{y} = -\mathbf{N}(\boldsymbol{\alpha})\ \mathbf{N}^T(\tilde{\boldsymbol{u}}^{flip})\ \hat{\boldsymbol{h}} \quad \text{ou} \quad \mathbf{N}(\boldsymbol{\alpha}^{prior})\ \boldsymbol{y} = -\mathbf{N}(\boldsymbol{\alpha})\ \mathbf{N}^T(\tilde{\boldsymbol{h}})\ \hat{\boldsymbol{u}}^{flip} \tag{18a, b}$$

Une comparaison des membres de gauche des équations (18b) (modèle convolutif) et (13c) (modèle ARX) suggère de prendre les mêmes valeurs pour les a_i et les α_i, c'est à dire $\boldsymbol{a}^{prior} = \boldsymbol{\alpha}^{prior}$. Dans ce cas, leurs membres de droite sont alors nécessairement égaux :

$$\mathbf{N}(\boldsymbol{u})\ \boldsymbol{b}^{prior} = -\mathbf{N}(\boldsymbol{a})\ \mathbf{N}^T(\tilde{\boldsymbol{h}})\ \hat{\boldsymbol{u}}^{flip} \text{ whith } \boldsymbol{a} \equiv \begin{bmatrix} a_1 & a_2 & \cdots & a_k \end{bmatrix}^T \tag{19}$$

3. Relations entre les coefficients *a* et *b* du modèle ARX

Une entrée test $u(t)$ est ici utilisée pour obtenir une expression explicite des b_j en fonction des a_i. Il s'agit ici d'une fonction créneau, voir (20a), où $\mathrm{H}(t)$ est la fonction de Heaviside et Q, soit une énergie absorbée par le système, en Joule, si l'entrée est une puissance thermique, soit une température, en Kelvin, si l'entrée est une variation de température. Les différentes matrices et vecteurs présents dans l'équation (19) sont facilement calculés:

$$u(t) = \frac{Q}{\Delta t}\left(\mathrm{H}(t) - \mathrm{H}(t-\Delta t)\right)\ ;\ \boldsymbol{u} = \tilde{\boldsymbol{u}} = \frac{Q}{\Delta t}\begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^T\ ;\ \mathbf{N}(\boldsymbol{u}) = \frac{Q}{\Delta t}\,\boldsymbol{I}\ ;\ \mathbf{N}^T(\tilde{\boldsymbol{h}})\ \hat{\boldsymbol{u}}^{flip} = Q\,\tilde{\boldsymbol{h}}^{flip} \tag{20a,b}$$

où $\boldsymbol{I}$ est la matrice identité. Après simplification, l'équation (19) devient :

$$\boldsymbol{b}^{prior} = -\mathbf{N}(\boldsymbol{a})\ \hat{\boldsymbol{h}}^{flip} \Rightarrow \boldsymbol{b}^{prior} = -\mathbf{N}(\hat{\boldsymbol{h}}^{flip})\ \boldsymbol{a} \tag{21a, b}$$

Cette équation illustre la relation qui lie les coefficients b aux coefficients a. Elle montre aussi que si toutes les observations sont prises en compte dans le modèle ARX, c'est-à-dire si $k = m$, on obtient $b_0 = -\tilde{h}_m \,\Delta t$. Ceci signifie que ce dernier coefficient n'est nul que si la durée de l'expérience d'identification est plus grande que le temps t_{nh} pour lequel $h\,(t)$ atteint un niveau nul avec une précision acceptable, qui dépend en fait du niveau du bruit de mesure dans une expérience physique, voir la section 2.1.

La valeur ci-dessus de $\boldsymbol{b}^{prior}$ est substituée dans l'équation (13c), voir (22a). Du fait des propriétés de commutativité et d'associativité du produit de convolution, cette équation s'écrit également sous la forme (22b):

$$\mathbf{N}(\boldsymbol{a}^{prior})\,\boldsymbol{y} = -\,\mathbf{N}(\boldsymbol{u})\,\mathbf{N}(\widehat{\boldsymbol{h}}^{flip})\,\boldsymbol{a} \quad ; \quad \mathbf{N}(\boldsymbol{a}^{prior})\,\boldsymbol{y} = -\,\mathbf{N}(\boldsymbol{a})\,\mathbf{N}(\widehat{\boldsymbol{h}}^{flip})\,\boldsymbol{u} \qquad (22a,b)$$

Le membre de gauche est décomposé en deux termes:

$$\mathbf{N}(\boldsymbol{a}^{prior})\,\boldsymbol{y} = \boldsymbol{y} + \mathbf{N}(\boldsymbol{a})\,\boldsymbol{y}^{prior} \quad \text{avec} \quad \boldsymbol{y}^{prior} = \begin{bmatrix} 0 & y_1 & y_2 & \cdots & y_{k-1} & y_k \end{bmatrix}^T \qquad (22c)$$

L'équation (22b) devient alors :

$$\boldsymbol{y} = -\,\mathbf{N}(\boldsymbol{a})\left(\boldsymbol{y}^{prior} + \mathbf{N}(\widehat{\boldsymbol{h}}^{flip})\,\boldsymbol{u}\right) \;\Rightarrow\; \boldsymbol{y} = \boldsymbol{X}\,\boldsymbol{a} \quad \text{avec} \quad \boldsymbol{X} = -\,\mathbf{N}\left(\boldsymbol{y}^{prior} + \mathbf{N}(\widehat{\boldsymbol{h}}^{flip})\,\boldsymbol{u}\right) \qquad (23)$$

Ceci permet d'estimer les paramètres a et b du modèle ARX, c'est à dire d'un modèle réduit issu du modèle convolutif et dont la réponse impulsionnelle exacte est connue sur une grille temporelle, par une expérience d'estimation de paramètres :

$$\hat{\boldsymbol{a}} = \boldsymbol{X}^{-1}\,\boldsymbol{y} \quad \text{and} \quad \hat{\boldsymbol{b}} = -\,\mathbf{N}(\widehat{\boldsymbol{h}}^{flip})\,\boldsymbol{X}^{-1}\,\boldsymbol{y} \qquad (23)$$

4. Conclusions

Nous avons montré ici qu'un modèle ARX constituait en fait une forme réduite d'un modèle convolutif. Pour que ses avantage dans son inversion, en termes de robustesse, et surtout de parcimonie, par rapport aux méthodes de déconvolution régularisée soient bien atteints, il faut écrire les deux dernières équations ci-dessus en réduisant le nombre de colonnes connexes de la matrice de sensibilité $\boldsymbol{X}$ pour atteindre les nombres optimaux de ses coefficients n_a, n_b et n_c, en ajoutant éventuellement un coefficient b_0 à ce modèle réduit.

Références

[1] W. Al Hadad and D. Maillet, Transfer function identification through total least squares, Journal of Physics: Conf. Series 1047 (2018), 012001, https://doi.org/10.1088/1742-6596/1047/1/012001

[2] W. Al Hadad, D. Maillet, and Y. Jannot, "Modeling unsteady diffusive and advective heat transfer for linear dynamical systems : A transfer function approach," International Journal of Heat and Mass Transfer, vol. 115, pp. 304–313, 2017.

[3] L. Ljung, System identication: theory for the user, 2nd ed. Upper Saddle River NJ: Prentice-Hall PTR, 1999.

[4] T. Loussouarn, D. Maillet, B. Schick, B. Rémy, D. Dan, Indirect measurement of temperature inside a furnace, ARX model identification, Journal of Physics: Conf. Series 1047 (2018), 012006, https://doi.org/10.1088/1742-6596/1047/1/012006

[5] A. François, L. Ibos, V. Feuillet, J. Meulemans, Estimation of the thermal resistance of a building wall with inverse techniques based on rapid active in situ measurements and white-box or ARX black-box models, Energy and Buildings, Volume 226, 1 November 2020, Article number 110346

[6] C. Zacharie, V. Schick, B. Rémy , G. Bergin , T. Mazet, R. Egal , Identification of Transfer Functions in a Vacuum Brazed Load with ARX Models, Instrumentation Mesure Métrologie, Vol.19, No.3, June, 2020, pp. 229-234, DOI: https://doi.org/10.18280/i2m.190308