POSSIBLE WAY OF IMPROVING THE QUALITY OF MODELLING FOR ADAPTIVE CONTROL

Tatiana Valentine Guy, Miroslav Kárný

Department of Adaptive Systems
Institute of Information Theory and Automation
Academy of Sciences of the Czech Republic

Abstract: This paper addresses the model development of a continuous time controlled system by using wavelet approximation of signals together with convolution technique of modelling. The method is expected to be of use for constructing adaptive controllers of LQG type using traditional autoregressive model. *Copyright © 2000 IFAC*

Keywords: Modelling, Filtering techniques, Adaptive control

1. INTRODUCTION

The theory of adaptive model-based controllers of Linear Quadratic Gaussian (LQG) type represents one of the basic topics of modern control theory (Åstrom and Wittenmark 1989). The design of these controllers is based upon a given quadratic index minimisation for a recursively identified model (Bobál *et al.* 1999). Autoregressive (AR) moving average (MA) model with external (X) inputs (ARMAX) is known to be adequate structure that is able to cope with correlated process noise. Its recursive identification cannot be, however, done exactly. For this reason, the simplified ARX model is employed. It often requires use of a long regression vector making the controller complex and not robust enough. This fact contributes to a relatively rare use of such controllers. Thus, a sort of data filtering is highly desirable.

The paper presents a novel promising solution relying on alternative modelling. It employs a convolution description of the continuous time controlled system and wavelet decomposition of involved signals.

The model is obtained as a combination of convolution modelling (Guy 1996) of a continuous time system and wavelet approximation of input/output signals. The approximation algorithm exploits multiresolution filtering (Mallat 1989*b*) of raw signals. The multiresolution approximations are calculated from real noisy measurements at a particular resolutions by employing filter banks (Vetterli and Herley 1992). The resulting sampled model is transformed into a regression type model with regressors formed by *multiresolution approximations* of input/output signals. The proposed approach is believed to simplify the design of the discrete adaptive controller of LQG type (Kárný *et al.* 1985).

The following assumptions are adopted in the work:

(A1) the controlled system is linear (linearisable) and time-invariant or slowly time-varying;
(A2) the relation of the system output to the system input and process noise is expressed by convolution operators;
(A3) the past history of the controlled process has only a limited-time effect on the future process behaviour.

For the simplicity, single-input single-output (SISO) stable systems are considered.

[1] This work was partially supported by the grants GA ČR 102/00/P045, 102/99/1292 and MŠMT ME 245/2000

2. PRELIMINARIES

This section introduces necessary notions, facts and relations used in the work.

2.1 *Convolution description of controlled system*

Generally, any affine, time-invariant continuous stochastic system relates output $y(\cdot)$, input $u(\cdot)$, external measurable disturbance $v(\cdot)$ and noise $\bar{e}(\cdot)$ signals. Signals are acted by unknown linear time-invariant, casual operators $\bar{\mathbf{A}}, \bar{\mathbf{B}}, \bar{\mathbf{D}}$ and $\mathbf{C}$, respectively:

$$\bar{\mathbf{A}}\, y(\cdot) + \bar{\mathbf{B}}\, u(\cdot) + \bar{\mathbf{D}}\, v(\cdot) + \mathbf{C}\, \bar{e}(\cdot) = 0. \quad (1)$$

Suppose, there is such an operator $\mathbf{C}^*$ that the transformed noise signal $e(\cdot) = \mathbf{C}^*\, \mathbf{C}\, \bar{e}(\cdot)$ becomes white discrete process after sampling with the shortest technically feasible period:

$$\mathbf{A}\, y(\cdot) + \mathbf{B}\, u(\cdot) + \mathbf{D}\, v(\cdot) + e(\cdot) = 0. \quad (2)$$

The transformed operators $\mathbf{A} = \mathbf{C}^*\, \bar{\mathbf{A}}$, $\mathbf{B} = \mathbf{C}^*\, \bar{\mathbf{B}}$ and $\mathbf{D} = \mathbf{C}^*\, \bar{\mathbf{D}}$ are still time invariant casual linear operators. The model (2) equivalent to (1) is a *continuous-time non-parametric description* of the controlled system we use here as a basic system model.

Using the convolution theorem, the operator $\mathbf{A}$ acting of the output signal $y(\cdot)$ at the time $t \in [0, T]$ can be represented in the integral form:

$$[\mathbf{A}\, y(\cdot)]_t = O_t^A + \int_0^t k_A(\tau) y(t-\tau)\, d\tau, \quad (3)$$

where function O_t^A reflects the initial conditions. The kernel $k_A(\tau)$ is considered to be a casual ($k_A(\tau) = 0$ for $\tau \leq 0$) smooth integrable function of a finite length.

By describing operators $\mathbf{B}$ and $\mathbf{D}$ similarly to $\mathbf{A}$ and by substituting the results in the basic system model (2), the convolution model gets form:

$$\int_0^t k_A(\tau) y(t-\tau) d\tau + \int_0^t k_B(\tau) u(t-\tau) d\tau + \int_0^t k_D(\tau) v(t-\tau) d\tau + e(\cdot) + O_t = 0, \quad (4)$$

where $O_t = O_t^A + O_t^B + O_t^D$ denotes the total offset reflecting the initial conditions. It tends to zero for the considered stable system.

2.2 *Multiresolution analysis*

The natural framework for the construction of wavelets is given by multiresolution analysis (Mallat 1989*a*) of $L^2(\mathbb{R})$ ($\forall f(t) \in L^2(\mathbb{R}) \Rightarrow \int |f(t)|^2 dt < \infty$) formed by a nested sequence of closed subspaces

$$\ldots \subset V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \ldots. \quad (5)$$

The subspaces are required to have the following properties:

(P1) $\bigcap_{j \in \mathbf{Z}} V = \emptyset$,

(P2) $\bigcup_{j \in \mathbf{Z}} V_j = L^2(\mathbb{R})$,

(P3) $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}$,

(P4) a basic scaling function $\phi \in V_0$ exists such that the set of $\phi_{0,k}(\cdot) = \phi(\cdot - k)$, $k \in \mathbf{Z}$ constitutes an orthogonal basis for V_0.

Using these properties, a two parametric array of functions is generated from basic scaling function $\phi(t)$

$$\phi_{j,k}(t) = 2^{j/2}\phi(2^j t - k), j, k \in \mathbf{Z}. \quad (6)$$

j stands for resolution parameter, k denotes a time translation index.

If for each $j \in \mathbf{Z}$ the orthogonal complement of V_j in V_{j+1} is defined as $W_j \bigoplus V_j = V_{j+1}$, then, any $L^2(\mathbb{R})$ function $f(t)$ can be written as a sum of the coarse approximation $P_{V_{J_0}}(t)$ at level $J_0 \in \mathbf{Z}$ and of details at higher levels

$$f(t) = P_{V_{J_0}}(t) + P_{W_{J_0}}(t) + P_{W_{J_0+1}}(t) + \ldots \quad (7)$$

In (7), P_{V_j} stands for the best approximation $f(t)$ in the approximating space V_j. P_{W_j} denotes the projection onto space W_j that covers the differences between the spaces P_{V_j} and $P_{V_{j+1}}$ spanned by various resolution scales of the scaling function. Similarly, as V_j is spanned by dilation and translations of the scaling function (6), W_j is spanned over the basis $\psi_{j,k}$ generated by translătions and dilation of the mother wavelet $\psi(t)$:

$$\psi_{j,k}(t) = 2^{j/2}\psi(2^j t - k), j, k \in \mathbf{Z}. \quad (8)$$

2.3 *Filter banks*

To calculate of the wavelet transform coefficients, a multirate filter bank (Vetterli and Herley 1992), producing sequences of coefficients at different rates is employed.

As any physical measuring device measures a real signal at a finite resolution and with a finite number of samples, there is an upper scale $j = J$, above which the signals details (P_{W_J}) are negligible small (Gopinath *et al.* 1994). Starting with a high resolution description of a signal, the analysis tree calculates the discrete wavelet transform (DWT) down to as low a resolution ($j = j_0$) as desired by having ($J - j_0$) stages. Fig.1 illustrates the analysis tree for three level decomposition.

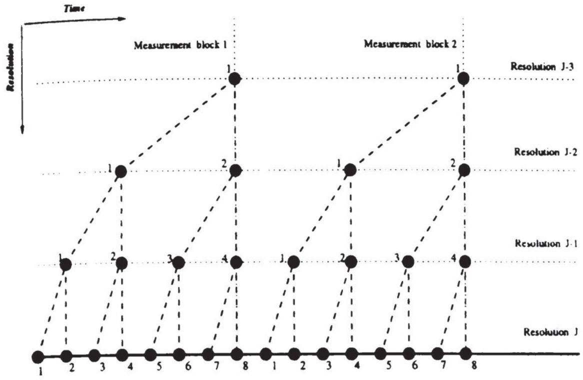

Fig. 1. Example of multiresolution measurement tree for two subsequent measurement blocks and four levels of resolution.

Therefore, for any signal $f(t) \in V_J$ a finite scale approximation in terms of dilated scaling functions and scaling-function coefficients $c_J(k)$ can be written as follows:

$$f(t) = \sum_k c_J(k)\phi_{J,k}(t). \tag{9}$$

Taking into account (7), the sum in (9) can be further decomposed in such a way

$$f(t) = \sum_k c_{j_0}(k)\phi_{j_0,k}(t) + \sum_k \sum_{j=j_0}^{J-1} d_j(k)\psi_{j,k}(t), \tag{10}$$

where $d_j(k)$ are scaling-function coefficients. The value j_0 sets the coarsest scale whose space is spanned by scaling functions $\phi_{j_0,k}(t)$. The rest of $L^2(\mathbb{R})$ is spanned by wavelets $\psi_{j,k}(t)$, providing the high resolution details of the signal. Hence, the first sum gives a coarse approximation of $f(t)$, while second term adds details (with increasing index j a higher resolution function is added).

Using low pass filter $h_L(n)$, derived from scaling function, together with high pass filter $h_H(n)$, gained from corresponding wavelet function, the wavelet decomposition is fully represented by sequence of scaling-function and wavelet coefficients

$$c_j(k) = \sum_n h_L(2n-k)c_{j+1}(k), \tag{11}$$

$$d_j(k) = \sum_n h_H(2n-k)d_{j+1}(k). \tag{12}$$

Scaling function coefficients (11) represent a lower rate resolution signal sequence, obtained by sampling of the low pass filter output and wavelet coefficients (12) provide additional details.

In practice, there is no need to deal with the scaling functions or wavelets directly. The only coefficients $c_{j_0}(k)$, $d_j(k)$ in (10) and scaling function filters h_L and h_H in (11), (12) are to be considered.

To perform a filtering, let us suppose a measurement block measured at the resolution level J (original data sequence) with the length of data block M^2 in the following form:

$$\mathcal{F}_k^J = [f(J, k-M+1), \ldots, f(J,k)]'. \tag{13}$$

A sequence a of the lower resolution signal can be obtained by low-pass filtering of $\mathcal{F}_k^J$ with subsequent downsampling of the output from low-pass filter (11).

The filter bank decomposition of the signal $f(t)$ into low-passed $\mathcal{F}_{k\mathbf{L}}^{J-1}$ and high-passed $\mathcal{F}_{k\mathbf{H}}^{J-1}$ sequences written in an operator form reads

$$\mathcal{F}_{k\mathbf{L}}^{J-1} = \mathbf{H_L}^{J-1}\ \mathcal{F}_k^J, \tag{14}$$

$$\mathcal{F}_{k\mathbf{H}}^{J-1} = \mathbf{H_H}^{J-1}\ \mathcal{F}_k^J, \tag{15}$$

where $\mathbf{H_L}$ and $\mathbf{H_L}$ are composed of low and high pass filters' responses. For Haar functions (Haar 1910), chosen in the work, they equal:

$$\mathbf{H_L} = [h_L(0)\ h_L(1)] = \left[\frac{1}{\sqrt{2}}\ \frac{1}{\sqrt{2}}\right],$$

$$\mathbf{H_H} = [h_H(0)\ h_H(1)] = \left[-\frac{1}{\sqrt{2}}\ \frac{1}{\sqrt{2}}\right]. \tag{16}$$

For example, the full three stage decomposition shown on Fig.1 is performed by the filter bank with the transformation matrix $\mathbf{T}^{J-3|J}$:

$$\begin{bmatrix} \mathcal{F}_{k\mathbf{L}}^{J-3} \\ \mathcal{F}_{k\mathbf{H}}^{J-3} \\ \mathcal{F}_{k\mathbf{H}}^{J-2} \\ \mathcal{F}_{k\mathbf{H}}^{J-1} \end{bmatrix} = \underbrace{\begin{bmatrix} \mathbf{H}_L{}^{J-3} & \mathbf{H}_L{}^{J-2} & \mathbf{H}_L{}^{J-1} \\ \mathbf{H}_H{}^{J-3} & \mathbf{H}_L{}^{J-2} & \mathbf{H}_L{}^{J-1} \\ & \mathbf{H}_H{}^{J-2} & \mathbf{H}_L{}^{J-1} \\ & & \mathbf{H}_H{}^{J-1} \end{bmatrix}}_{\mathbf{T}^{J-3|J}} \mathcal{F}_k^J. \tag{17}$$

In order to filter the noise corrupting the signals, outputs from low pass filter $\mathcal{F}_{k\mathbf{L}}^{J-1}$ (14) are only used. Generally, the transformation matrix making low-pass filtering up to the desired level of resolution j_0 is given by:

$$\mathbf{T}_\mathbf{L}^{j_0|J} = \prod_{l=1}^{J-j_0} diag\{\underbrace{\mathbf{H_L}, \ldots, \mathbf{H_L}}_{2^{J-j_0-l}}\}. \tag{18}$$

3. APPROXIMATE MODEL

This section focuses on the building of the approximate system model using the multiresolution approximations of the input-output signals. The real time implementation of the multiresolution approximation algorithm is performed by filter banks. It employs the measurement blocks containing a predefined finite number of measurements (see footnote [2]). The measurement block

[2] The data block length should satisfy the requirement $M = 2^{J-j_0}$, where j_0 denotes the desired level of resolution.

(13) within the range given by control period is chosen. It can be done as the considered control period is greater than the measurement period.

By replacing input/output signals with their multiresulotion approximations (9), (4) converts to the form:

$$\int_0^t k_A(t-\tau)\sum_k c_J^y(k)\phi_{J,k}^y(\tau)d\tau + \int_0^t k_B(t-\tau)\sum_k c_J^u(k)\phi_{J,k}^u(\tau)d\tau + \int_0^t k_D(t-\tau)\sum_k c_J^v(k)\phi_{J,k}^v(\tau)d\tau + e(t) = 0. \tag{19}$$

Starting with a high resolution description of a signal, the DWT is calculated down to a desired resolution j_0 (see (10)). In order to filter the noise, the details (second term in (10)) are omitted. Then, (19) reads

$$\sum_k \left\{ c_{j_0}^y(k)\int_0^t k_A(t-\tau)\phi_{j_0,k}^y(\tau)d\tau + c_{j_0}^u(k)\int_0^t k_B(t-\tau)\phi_{j_0,k}^u(\tau)d\tau + c_{j_0}^v(k)\int_0^t k_D(t-\tau)\phi_{j_0,k}^v(\tau)d\tau \right\} + e(t) = 0 \tag{20}$$

with $c_{j_0}^{\bullet}(k)$ $(\bullet = \{y,u,v\})$ are scaling-function coefficients at the desired level of resolution j_0.

The values $c_{j_0}^{\bullet}(k)$ are computed by filter bank (see (14)) as outputs of low-pass filter. Hence, (20) can be written

$$\sum_k \left\{ y_{\mathbf{L}}^{j_0}(k)\int_0^t k_A(t-\tau)\phi_{j_0,k}^y(\tau)d\tau + u_{\mathbf{L}}^{j_0}(k)\int_0^t k_B(t-\tau)\phi_{j_0,k}^u(\tau)d\tau + v_{\mathbf{L}}^{j_0}(k)\int_0^t k_D(t-\tau)\phi_{j_0,k}^v(\tau)d\tau \right\} + e(t) = 0, \tag{21}$$

where $y_{\mathbf{L}}^{j_0}(k)$, are *discrete approximations of the output signal at the resolution level* j_0 obtained by

$$y_{\mathbf{L}}^{j_0}(k) = \mathbf{T}_{\mathbf{L}}^{J-j_0}\ \mathcal{Y}_k^j. \tag{22}$$

$\mathbf{T}_{\mathbf{L}}^{J-j_0}$ stands for the low-pass part of transformation matrix (see (18)) for the filter bank decomposition and $\mathcal{Y}_k^j$ denotes a kth measurement block for the output signal $y(t)$.

The approximations of the input $u_{\mathbf{L}}^{j_0}(k)$ and measurable disturbance $v_{\mathbf{L}}^{j_0}(k)$ signals are made in the same way.

Let us define the following coefficients for $k \in \mathbf{Z}$

$$a_k(t) = \int_0^{\min\{t,L_A\}} k_A(t-\tau)\phi_{j_0,k}^y(\tau)d\tau$$
$$b_k(t) = \int_0^{\min\{t,L_B\}} k_B(t-\tau)\phi_{j_0,k}^u(\tau)d\tau$$
$$d_k(t) = \int_0^{\min\{t,L_D\}} k_D(t-\tau)\phi_{j_0,k}^v(\tau)d\tau. \tag{23}$$

The majority of scaling functions have finite support or they are at least fast decaying. This fact together with the assumption (A3) ensure that the number of non-zero coefficients, defined by (23) for a fixed t, is finite. This number is determined by lengths L_A, L_B and L_D of kernel supports and by width of the used scaling function. Consequently, the only finite summation over k in (21) is required.

The model (21) under the notations (23) can be rewritten as follows

$$\sum_{k=0}^{m_A-1} a_k(t)\, y_{\mathbf{L}}^{j_0}(k) + \sum_{k=0}^{m_B-1} b_k(t)\, u_{\mathbf{L}}^{j_0}(k) + \sum_{k=0}^{m_D-1} d_k(t)\, v_{\mathbf{L}}^{j_0}(k) + e(t) = 0, \tag{24}$$

where m_A, m_B, m_D denote finite numbers of non-zero coefficients $a_k(t)$, $b_k(t)$ and $d_k(t)$, respectively.

The sampling moments t_i are considered that make the coefficients (23) time invariant and ensure that the sampled noise $\{e_i\}_{i=0}^{\infty}$ is white discrete-time process with $e_i = e(t_i)$.

Let us order the filtered data into the data vector:

$$z_k = [y_{\mathbf{L}}^{j_0}(k), \ldots, y_{\mathbf{L}}^{j_0}(k-m_A), u_{\mathbf{L}}^{j_0}(k), \ldots, u_{\mathbf{L}}^{j_0}(k-m_B), v_{\mathbf{L}}^{j_0}(k), \ldots]', \tag{25}$$

and the corresponding unknown coefficients into the vector of unknown parameters:

$$\Theta = [a_0, \ldots, a_{m_A}, b_0, \ldots, b_{m_B}, d_0, \ldots, d_{m_D}]'. \tag{26}$$

Then, (24) reads

$$\Theta'\, z_k + e_k = 0, \tag{27}$$

which is the linear regression type model with regressor vector (25) composed of *filtered* values of involved signals – their multiresolution approximations.

4. ALGORITHMIC DESCRIPTION

The highest level of this measurement tree (resolution level J) corresponds to real measurements. At each sampling rate, two subsequent measurements at the resolution level J are used to compute the lower rate signal approximations at the resolution level $J-1$. The computed approximations at level

$J-1$ are, then, used to find a value of approximations at the level $J-2$. Similarly, value of approximation at the coarsest level is computed from measurements gained at the level $J-2$ (see Fig.1).

Thus, the highest rate data sequence is the original data sequence, measured on a process, while the lower rate levels of the multiresolution analysis tree contain data sequences with reduced noise (see Fig.1). As each lower rate approximation represents several full rate measurements, it also preserves the vital information about the original signal.

The data are processed by measurement blocks. Each of them contains signal measurements between two control points. At each control point, the data block preceding to the point is processed. Then, the signal approximation at the highest resolution level is used in the system model (4). In this way, the required discretisation of the model (4) is reached and it converts into the ARX model. Estimates of model coefficients are updated and used in control design (Guy 1999).

The evaluations can be summarised into the following steps:

Step 1 The measurement block is acquired within single control period and filtered using (22).

Step 2 The multiresolution approximation of the signals gained at Step 1 are used for identification of ARX model (27) and subsequent control design.

Step 3 The output of control design procedure is generated in the form of (22) and can be reconstructed and applied over the next control period.

5. CONCLUDING REMARKS

A novel modelling of continuous time-system for digital adaptive control was presented. It combines a convolution description of the system with wavelet-based approximation of involved noisy signals. Under realistic assumptions it leads to the celebrated and matured ARX model working on coefficients arising in wavelet transformation. The discrete wavelet transformation is implemented by filter bank that suppresses corrupting details of processed signals.

An extensive use of the resulting algorithm calls for solution of several practical subtasks. The key one seems to be the choice of an appropriate resolution level j_0. It should be high enough so that no significant information about involved signals is lost. It should be small enough so that undesired noises are removed from the signals.

Even under the current state of knowledge, the resulting algorithm is applicable and offers the following advantages

- measurements are filtered and control design runs on filtered data unlike in direct use of ARX models,
- length of regression vectors and, consequently, computational complexity decrease comparing to ARX models that tries to cope with coloured noise that would require ARMAX model,
- a sort of desirable filtering is induced by the use of wavelet approximation into the quadratic performance index,
- (prior) control design can be run at different resolution levels so that optimal filtering can be adapted directly to the modelling aim.

6. REFERENCES

Åstrom, K.J. and B. Wittenmark (1989). *Adaptive Control.* Addison-Wesley Publishing. Reading, Massachusetts.

Bobál, V., J. Böhm and P. Prokop (1999). Practical aspects of self-tuning controllers. *International Journal of Adaptive Control and Signal Processing* **13**(8), 671–690.

Gopinath, R.A., J.E. Odegard and C.S. Burrus (1994). Optimal wavelet representation of signals and the wavelet sampling theorem. *IEEE Transactions on Circuits and Systems II* **41**(4), 262–277.

Guy, T.V. (1996). Spline-based modelling of controlled system. In: *Preprints of the 2nd European IEEE Workshop CMP'96.* Prague, Czech Republic. pp. 73–78.

Guy, T.V. (1999). Hybrid Adaptive Controller. PhD thesis. Czech Technical University. Prague, Czech Republic.

Haar, A. (1910). Zur Theorie der orthogonalen Funktionensysteme. *Mathematische Annalen* **69**, 331–371.

Kárný, M., A. Halousková, J. Böhm, R. Kulhavý and P. Nedoma (1985). Design of linear quadratic adaptive control: Theory and algorithms for practice. *Kybernetika.*

Mallat, S.G. (1989*a*). Multiresolution approximation and wavelet orthonormal bases of L^2. *Transactions of the American Mathematic Society* **315**, 69–87.

Mallat, S.G. (1989*b*). A theory of multiresolution decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**(7), 674–693.

Vetterli, M. and C. Herley (1992). Wavelets and filter banks: theory and design. *IEEE Transactions on Acoustics, Speech and Signal Processing* **40**(9), 2207–2232.