Control and Cybernetics

vol. **31** (2002) No. 1

Simple stable discrete-time generalised predictive control with anticipated filtration of control error

by

Zdzisław Kowalczuk and Piotr Suchomski

Department of Automatic Control,
Faculty of Electronics, Telecommunication and Computer Science,
Technical University of Gdańsk,
Narutowicza 11/12, 80-952 Gdańsk, Poland,
E-mail: kova@pg.gda.pl

Abstract: It is shown that under some specific conditions, the solution of the generalised predictive control (GPC) design using the concept of anticipated filtering (AF) of the control error always exists, and that such a design leads to stable control systems with definite closed-loop characteristics. The plant cancellation issue is taken into account, and it is demonstrated that certain bounds on GPC design parameters have to be considered. An iterative procedure for simultaneous determination of the three basic design-tuning parameters: the control horizon, the controller gain, and the order of plant cancellation, is also supplied. An important feature of this approach is that the anticipated filtering makes it possible to reduce a disagreeable control effort associated with GPC and to make the λ-tuning mechanism practicable. The bounds on the GPC design parameters are discussed, and certain optimal tuning rules are proposed and validated via simulated experiments.

Keywords: discrete-time systems, system design, non-minimal systems, predictive control.

1. Introduction

The predictive control strategy is very well matched to modern control system design procedures. The mechanism of additional filtering of the control error has been proposed for the generalised predictive control design in (Clarke et al., 1987, Clarke and Mohtadi, 1987, Demircioglu and Gawthrop, 1991). The purpose of this mechanism is to fictitiously abate the excitation of the closed-loop control system. The filtration is performed in the anticipation-time domain and is referred to as the anticipated filtering (AF). This approach can be successfully exercised in both discrete-time and continuous-time domains (Kowalczuk and Suchomski, 1995, 1996, 1999a, 1999b, Kowalczuk, et al., 1996, Zhang, 1996).

In this paper it is shown that with the AF approach, under certain conditions including the cancellation issue, the solution of the GPC design always exists and the design leads to stable control systems with definite closed-loop characteristics. Moreover, it is demonstrated that some bounds on the GPC design parameters have to be taken into account. An iterative procedure is also proposed which solves the problem of determining two design quantities (the control horizon and the controller gain), without the prior knowledge of the order of plant cancellation. An important feature of this approach is that the anticipated filtering makes it possible to reduce the disagreeable control effort associated with GPC and to make the λ-tuning mechanism more suitable for implementation.

2. AF-GPC design principles

Consider a CARIMA model of a linear system (Clarke et al., 1987)

$$A(q^{-1})y(t) = B(q^{-1})u(t) + \Delta^{-1}C(q^{-1})e(t) \tag{1}$$

where t is the discrete-time index, $\{u(t)\}$ and $\{y(t)\}$ are the input and output of the controlled system, respectively, $\{e(t)\}$ is a zero-mean white-noise disturbance, q^{-1} is the backward shift operator, $\Delta = 1-q^{-1}$ is the difference operator, and

$$\begin{aligned} A(q^{-1}) &= 1 + a_1 q^{-1} + \cdots + a_{N_A} q^{-N_A}, \\ B(q^{-1}) &= b_1 q^{-1} + \cdots + b_{N_B} q^{-N_B}, \\ C(q^{-1}) &= 1 + c_1 q^{-1} + \cdots + c_{N_C} q^{-N_C} \end{aligned} \tag{2}$$

Let

$$J(\Delta \boldsymbol{u}) = \mathrm{E}\Big\{ \sum_{i=N_1}^{N_2} \Big[w(t+i) - y(t+i) \Big]^2 + \lambda \sum_{i=1}^{N_u} \Delta u(t+i-1)^2 \Big\} \tag{3}$$

be a cost function, where $\{w(t)\}$ denotes a reference sequence, N_1 and N_2 are the bottom and top of the observation horizon, respectively, N_u is the length of the control horizon, $\lambda \geq 0$ is the control weighting factor, E is the expectation operator conditioned on data up to t, and

$$\Delta \boldsymbol{u}(t) = \begin{bmatrix} \Delta u(t) & \cdots & \Delta u(t+N_u-1) \end{bmatrix}^{\mathrm{T}}, \quad \Delta \boldsymbol{u}(t) \in \Re^{N_u}$$

denotes the incremental control sequence searched for. To facilitate further discussion, let us introduce also an auxiliary notion of the effective observation horizon

$$N_0 = N_2 - N_1 + 1. \tag{4}$$

The optimal, in the minimum variance sense, i-step ahead predictor of y (Clarke and Mohtadi, 1987, Favier, 1987) is given by

$$\widehat{y}(t+i) = H_i(q^{-1})\Delta u(t+i-1) + \widehat{y}(t+i|t), \tag{5}$$

where $\widehat{y}(t+i|t)$ satisfies the following equation

$$C(q^{-1})\widehat{y}(t+i|t) = E_i(q^{-1})\Delta u(t-1) + G_i(q^{-1})y(t), \qquad i = 1,\dots,N_2 \tag{6}$$

The polynomials H_i, G_i and E_i, $i = 1,\dots,N_2$, can be obtained from the Diophantine equations (Clarke and Mohtadi, 1987, Favier, 1987, Gorez, et al., 1987)

$$\widehat{A}(q^{-1})F_i(q^{-1}) + q^{-i}G_i(q^{-1}) = C(q^{-1}), \tag{7}$$

$$C(q^{-1})H_i(q^{-1}) + q^{-i}E_i(q^{-1}) = \bar{B}(q^{-1})F_i(q^{-1}) \tag{8}$$

where

$$\widehat{A}(q^{-1}) = \Delta A(q^{-1}), \quad \bar{B}(q^{-1}) = qB(q^{-1}) \tag{9}$$

$$\deg E_i(q^{-1}) = \max(N_B - 2, N_C - 1), \ \deg F_i(q^{-1}) = i - 1,$$

$$\deg G_i(q^{-1}) = N_A \text{ and } \deg H_i(q^{-1}) = i - 1.$$

Assuming that $\Delta u(t+i-1) = 0$ for $i > N_u$, (if it exists) the optimal control sequence $\Delta \boldsymbol{u}^*$ minimising (3) takes the following form (Clarke et al., 1987, Clarke and Mohtadi, 1987):

$$\Delta \boldsymbol{u}^*(t) = \boldsymbol{K}(\boldsymbol{w}(t) - \widehat{\boldsymbol{y}}(t|t)) \tag{10}$$

where

$$\boldsymbol{K} = (\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H} + \lambda \boldsymbol{I})^{-1}\boldsymbol{H}^{\mathrm{T}}\boldsymbol{K} \in \Re^{N_u \times N_0} \tag{11}$$

$$\boldsymbol{H} = \begin{bmatrix} h_{N_1-1} & h_{N_1-2} & \cdots & h_{N_1-N_u} \\ h_{N_1} & h_{N_1-1} & \cdots & h_{N_1-N_u+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_2-1} & h_{N_2-2} & \cdots & h_{N_2-N_u} \end{bmatrix}, \quad \boldsymbol{H} \in \Re^{N_0 \times N_u}, \tag{12}$$

with $h_k = 0$ for $k < 0$ and

$$\boldsymbol{w}(t) = \begin{bmatrix} w(t+N_1) & \cdots & w(t+N_2) \end{bmatrix}^{\mathrm{T}}, \quad \boldsymbol{w}(t) \in \Re^{N_0}$$

$$\widehat{\boldsymbol{y}}(t|t) = \begin{bmatrix} \widehat{y}(t+N_1|t) & \cdots & \widehat{y}(t+N_2|t) \end{bmatrix}^{\mathrm{T}}, \quad \widehat{\boldsymbol{y}}(t|t) \in \Re^{N_0}.$$

By picking up the first row of $\boldsymbol{K}$

$$\boldsymbol{k}^{\mathrm{T}} = \begin{bmatrix} k_1 & \cdots & k_{N_0} \end{bmatrix} \tag{13}$$

a practical GPC control law is obtained given by the following 'single-sample' formula

$$\Delta u^*(t) = \boldsymbol{k}^{\mathrm{T}}(\boldsymbol{w}(t) - \widehat{\boldsymbol{y}}(t|t)) \tag{14}$$

in which only the first element of $\Delta \boldsymbol{u}^*(t)$ is used as the control input. In order to judge certain cardinal consequences of such a settlement, and taking into account the fact that only the first sample of the calculated control sequence is applied, the notion of a relative range of realisation (RRR) of the control sequence is introduced as

$$\mathrm{RRR} = 1/N_u \tag{15}$$

Assuming that within the observation horizon the future set point is known and constant, that is $w(t+i) = w(t)$ for $i = 1, \dots, N_2$, we can define a future incremental reference $\delta\bar{w}(\hat{t})$, equivalent to a filtered error sequence $\widehat{e}(\hat{t})|_{\hat{t}=t+i} = r_i e(t)$, based on the concept of anticipated filtering (Kowalczuk and Suchomski, 1995, 1996, 1999):

$$\delta\bar{w}(\hat{t})|_{\hat{t}=t+i} = \widehat{e}(t+i) = r_i e(t) = r_i(w(t) - y(t)), \quad i = 1, \dots, N_2 \tag{16}$$

where r_i, $i = 1, \dots, N_2$, are the coefficients of the step response of an instrumental filter used in such anticipation. The above sequence can than be used as a reference trajectory for the predicted incremental plant output

$$\delta\widehat{y}(t+i) = \widehat{y}(t+i) - y(t)$$

If, in view of the above arrangements, the objective of the design is to drive the signal $\delta\widehat{y}(t+i)$ to the reference signal $\delta\bar{w}(t+i)$, $i = 1, \dots, N_2$, the following modified cost function has to be considered

$$\widehat{J}(\Delta \boldsymbol{u}(t)) = \sum_{i=N_1}^{N_2} [\delta\bar{w}(t+i) - \delta\widehat{y}(t+i)]^2 + \lambda \sum_{i=1}^{N_u} \Delta u(t+i-1)^2 \tag{17}$$

Minimisation of the above criterion yields the following incremental control action

$$\Delta u^*(t) = \boldsymbol{k}^{\mathrm{T}}(\delta\bar{\boldsymbol{u}}(t|t) - \delta\widehat{\boldsymbol{y}}(t|t)) \tag{18}$$

where the vectors from $\Re^{N_0}$

$$\delta\bar{\boldsymbol{u}}(t|t) = (w(t) - y(t))\boldsymbol{r} = (w(t) - y(t))[\ r_{N_1} \ \cdots \ r_{N_2}\ ]^{\mathrm{T}} \tag{19}$$
$$\delta\widehat{\boldsymbol{y}}(t|t) = [\ \widehat{y}(t+N_1|t) - y(t) \ \cdots \ \widehat{y}(t+N_2|t) - y(t)\ ]^{\mathrm{T}}.$$

By virtue of (6) it can be shown that the incremental control law (18) is expressed through

$$\begin{aligned} &C(q^{-1})\Delta u^*(t) \\ &= gC(q^{-1})(w(t) - y(t)) - L(q^{-1})\Delta u^*(t) - M(q^{-1})y(t) \end{aligned} \tag{20}$$

where

$$g = \sum_{i=1}^{N_2-N_1+1} k_i r_{N_1+i-1}, \tag{21}$$

$$L(q^{-1}) = q^{-1} \sum_{i=1}^{N_2-N_1+1} k_i E_{N_1+i-1}(q^{-1}), \tag{22}$$

$$M(q^{-1}) = \sum_{i=1}^{N_2-N_1+1} k_i (G_{N_1+i-1}(q^{-1}) - C(q^{-1})) \tag{23}$$

with $\deg L(q^{-1}) = \max(N_B - 1, N_C)$, and $\deg M(q^{-1}) = \max(N_A, N_C)$.

It follows from (20) that the closed-loop characteristic polynomial $D(q^{-1})$ of the resultant GPC control system with the 'nominal' plant model (1) is given by

$$D(q^{-1}) = \widehat{A}(q^{-1})C(q^{-1}) + gB(q^{-1})C(q^{-1}) + \widehat{A}(q^{-1})L(q^{-1}) \\ +B(q^{-1})M(q^{-1}) \tag{24}$$

By virtue of (7) and (8) it can be shown that

$$D(q^{-1}) = C(q^{-1})\widetilde{D}(q^{-1}) \tag{25}$$

where

$$\begin{aligned} \widetilde{D}(q^{-1}) &= \widehat{A}(q^{-1}) - q^{-1}\widetilde{A}(q^{-1}) + g^{*}B(q^{-1}) \\ g^{*} &= g - \sum_{i=1}^{N_0} k_i = \sum_{i=1}^{N_2-N_1+1} k_i(r_{N_1+i-1} - 1), \\ \widetilde{A}(q^{-1}) &= \sum_{i=1}^{N_2-N_1+1} k_i(\widehat{A}(q^{-1})H_{N_1+i-1}(q^{-1}) - \bar{B}(q^{-1}))q^{N_1+i-1}. \end{aligned} \tag{26}$$

Since $C(q^{-1})$ is assumed to be stable, the closed loop system is stable if and only if $\widetilde{D}(q^{-1})$ is stable.

3. Properties of the AF-GPC design

In spite of a general usefulness of the idea of control weighting in the cost formulae of (3) and (17), the analytical properties of the resulting control system designed with $\lambda > 0$ are not explicit. On the other hand, by putting $\lambda = 0$ and taking into consideration the effect of the anticipation filter (instead) one is able to resolve explicitly for the analytical properties of the GPC control system without losing sight of the control effort.

After letting $\lambda = 0$, two cases with respect to the reducibility attribute of polynomials $\widehat{A}(q^{-1})$ and $B(q^{-1})$ have to be considered. The first case concerns irreducible (relatively prime) polynomials and the second one treats the polynomials that have a common factor. To show the relationship existing between $\widehat{A}(q^{-1})$, $\bar{B}(q^{-1})$ of (9) and $\{h_k s\}$, being the sequence of Markov parameters of the open-loop system $\bar{B}(q^{-1})/\widehat{A}(q^{-1})$, it is helpful to bring up the following double infinite lower-triangular Toeplitz matrix, in which $\boldsymbol{H}$ can be identified as a submatrix

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{N_B} \\ 0 \\ \vdots \end{bmatrix} = \begin{bmatrix} h_0 & 0 & 0 & \cdots \\ h_1 & h_0 & 0 & \cdots \\ h_2 & h_1 & h_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} 1 \\ \widehat{a}_1 \\ \vdots \\ \widehat{a}_{N_A+1} \\ 0 \\ \vdots \end{bmatrix} \tag{27}$$

It can be shown (Kowalczuk and Suchomski, 1995, 1996) that the anticipated filtering approach to the GPC design can lead to stable control systems with a desired corresponding closed-loop pole placement (for $\lambda = 0$) and that under certain conditions the GPC design solution (11) always exists. The fundamentals of this result are given below in the form of three theorems.

3.1. Relatively Prime Polynomials $\widehat{A}(q^{-1})$ & $B(q^{-1})$

THEOREM 1 *For $\lambda = 0$, if $\widehat{A}(q^{-1})$ and $B(q^{-1})$ are relatively prime and the following conditions*

(c1) $N_2 \geq N_1 + N_u - 1$

(c2) $\operatorname{rank} \boldsymbol{H} = N_u$

(c3) $N_1 \geq N_B$

(c4) $N_u \geq N_A + 1$

are satisfied, then the closed-loop characteristic polynomial $D(q^{-1})$ is determined by

$$D(q^{-1}) = C(q^{-1})\widetilde{D}(q^{-1}) \tag{28}$$

where

$$\widetilde{D}(q^{-1}) = 1 + g^* B(q^{-1}) \quad \blacksquare \tag{29}$$

Proof. Firstly, let us note that for $\lambda = 0$ from (c1) and (c2) we have $\boldsymbol{K} = (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T$ and $\boldsymbol{K}\boldsymbol{H} = \boldsymbol{I}$, where $\boldsymbol{I} \in \Re^{N_u \times N_u}$ is the identity matrix. Thus the entries of $\boldsymbol{k}^T\boldsymbol{H}$ can be expressed as

$$\sum_{i=1}^{N_2-N_1+1} k_i h_{N_1+i-k-1} = \begin{cases} 1 & \text{for } k = 1, \\ 0 & \text{for } k = 2, \ldots, N_u. \end{cases} \tag{30}$$

By using (26), (27) and (c3) the polynomial $\widetilde{A}(q^{-1})$ can be expressed as

$$\widetilde{A}(q^{-1}) = \sum_{j=0}^{N_A} v_j q^{-j} \tag{31}$$

where

$$v_j = \sum_{k=1}^{N_A-j+1} \widehat{a}_{k+j}\tau_k, \qquad j = 0,\ldots,N_A \tag{32}$$

and

$$\tau_k = \sum_{i=1}^{N_2-N_1+1} k_i h_{N_1+i-k-1}, \qquad k = 1,\ldots,N_A+1 \tag{33}$$

From (30)–(33) and (c4) it results that

$$v_j = \widehat{a}_{j+1}, j = 0,\ldots,N_A \tag{34}$$

and consequently

$$\widetilde{A}(q^{-1}) = \sum_{j=0}^{N_A} \widehat{a}_{j+1} q^{-j} \tag{35}$$

Now, from (49) it may be concluded that

$$\widehat{A}(q^{-1}) - q^{-1}\widetilde{A}(q^{-1}) = 1. \tag{36}$$

That means that $\widetilde{D}(q^{-1}) = 1 + g^* B(q^{-1})$, what concludes the proof. ■

REMARK 1 *Note that for* $r_i = 1$, $i = N_1,\ldots,N_2$, *the closed-loop characteristic polynomial* $D(q^{-1})$ *in (26) is completely determined by the observer polynomial* $C(q^{-1})$

$$D(q^{-1}) = C(q^{-1}) \quad ■ \tag{37}$$

REMARK 2 *As it can easily be seen from (12), (27) and (c1), if* $N_u > N_A + 1$ *and* $N_1 > N_B$ *then it is clear that*

$$\boldsymbol{H}[\,1 \quad \widehat{a}_1 \quad \cdots \quad \widehat{a}_{N_A+1} \quad 0 \quad \cdots\,]^{\mathrm{T}} = \boldsymbol{0} \tag{38}$$

and, consequently, $\boldsymbol{H}$ *cannot have full column rank* ($\operatorname{rank}\boldsymbol{H} < N_u$). *Thus, it can be inferred that certain upper bounds on the design parameters must be observed in order to assure the full column rank of* $\boldsymbol{H}$ *and the existence of (11).* ■

LEMMA 1 *Let us assume that one of the following set of conditions is satisfied*

(C1′)	$N_2 \geq N_1 + N_A$		(C1″)	$N_2 \geq N_1 + N_u - 1$
(C2′)	$N_1 > N_B$	*or*	(C2″)	$N_1 = N_B$
(C3′)	$N_u \leq N_A + 1$		(C3″)	$N_u \geq N_A + 1$

then $\boldsymbol{H}$ *has full column rank if and only if* $\widehat{A}(q^{-1})$ *and* $B(q^{-1})$ *are relatively prime.* ■

Proof. In the following draft proof we consider only the first set of conditions as the most crucial one for our discussion. It is thus enough to consider the limiting case of $N_u = N_A + 1$. Keeping in mind that systems with different zero-pole cancellations have the same Markov parameters, it follows from (12) and (27) that

$$\boldsymbol{H}\begin{bmatrix} 1 \\ \widehat{a}_1 \\ \vdots \\ \widehat{a}_{N_A} \end{bmatrix} = -\widehat{a}_{N_A+1}\begin{bmatrix} h_{N_1-N_u-1} \\ h_{N_1-N_u} \\ \vdots \\ h_{N_2-N_u-1} \end{bmatrix} \tag{39}$$

On the one hand, if the polynomials $\widehat{A}(q^{-1})$ and $B(q^{-1})$ are reducible then there exists a reduced polynomial $\widehat{A}^0(q^{-1})$ having coefficients $\widehat{a}_i^0$, $i = 0, \ldots, N_A^0$, where $N_A^0 < N_A$. As $\widehat{a}_{N_A+1} = 0$, the right hand side of (39) is zero and the matrix $\boldsymbol{H}$ cannot have full rank. Hence, by using contradiction, one can state that if $\boldsymbol{H}$ has full column rank then the polynomials $\widehat{A}(q^{-1})$ and $B(q^{-1})$ are relatively prime.

On the other hand, it results from (39) that if $\operatorname{rank}\boldsymbol{H} < N_A + 1$ then there must be a reduced-order system having the same Markov parameters as the original system (1). Therefore there must exist a different polynomial $\widehat{A}^0(q^{-1})$ of a lower degree that fulfils the 'matching' relation similar to (39). This concludes the proof by contradiction that if the polynomials $\widehat{A}(q^{-1})$ and $B(q^{-1})$ are relatively prime then matrix $\boldsymbol{H}$ has full column rank. ■

We can now present our principal result.

THEOREM 2 *If $\widehat{A}(q^{-1})$ and $B(q^{-1})$ are relatively prime and one of the following set of conditions*

(C1′)	$N_2 \geq N_1 + N_u - 1$		(C1″)	$N_2 \geq N_1 + N_u - 1$
(C2′)	$N_1 > N_B$	*or*	(C2″)	$N_1 = N_B$
(C3′)	$N_u = N_A + 1$		(C3″)	$N_u \geq N_A + 1$

is satisfied, then $\boldsymbol{H}$ has full column rank ($\operatorname{rank}\boldsymbol{H} = N_u$), *the solution (11) exists, and, for $\lambda = 0$, the closed-loop characteristic polynomial $D(q_{-1})$ is determined by*

$$D(q^{-1}) = C(q^{-1})\widetilde{D}(q^{-1}) \tag{40}$$

where

$$\widetilde{D}(q^{-1}) = 1 + g^* B(q^{-1}) \quad ■ \tag{41}$$

REMARK 3 *For $r_i = 1$, $i = N_1, \ldots, N_2$ in (26) the characteristic polynomial $D(q^{-1})$ is completely determined by the observer polynomial $C(q^{-1})$*

$$D(q^{-1}) = C(q^{-1}) \quad ■ \tag{42}$$

REMARK 4 *Contrary to a popular rule of thumb used in tuning the GPC controllers, setting N_1 to the value of the plant's transportation delay, say κ, does not guarantee solvability of the design problem (11) because Theorem 2 states that (with $\lambda = 0$) N_1 cannot be lower than the degree of the numerator N_B. Note that by using this advice, apart from skipping the first κ samples, one also rejects additional $N_B - \kappa$ seemingly informative samples of the output signal. Consequently, N_B can be interpreted as a lower "information boundary" recognised in the system's output sequence that allows for isolation of the piece of system's information that is most essential from the prediction viewpoint.* ■

REMARK 5 *With reference to the prime set of the triple condition (Ci') it is now evident that in order to assure the main result of Theorem 2 one has to put $N_u = N_A + 1$. In the case of the parsimonious choice of $N_1 = N_B$ (see the secondary triple - Ci") the upper bound does not exist.* ■

REMARK 6 *The first condition in both sets (C1' and C1") represents the necessary 'geometric' prerequisite $N_u \leq N_0$ that says that the effective observation horizon $N_0 = N_2 - N_1 + 1$ should not be shorter than the assumed length N_u of the designed control sequence $\Delta \boldsymbol{u}(t)$ (see eqns. 10 through 12).* ■

REMARK 7 *The above limitation combined with conditions C3' and C3" leads to a bilateral restriction on N_u:*

$$N_A + 1 \leq N_u \leq N_0 = N_2 - N_1 + 1 \tag{43}$$

Hence it results that the GPC observation horizon has also a limit: $N_0 \geq N_A + 1$. Since matrix $\boldsymbol{H}$ is $N_0 \times N_u$ and $\operatorname{rank} \boldsymbol{H} = N_u$ must be satisfied, N_u is the critical parameter sought after. ■

REMARK 8 *Note that $N_u \geq N_A + 1$ fulfils both of the C3 conditions at the same time. Thus, taking into account that $N_1 \geq N_B$ can be accepted as the second restriction C2, a single common set of conditions can be specified as follows:*

$$\begin{aligned} &(\mathrm{C1}) \quad N_2 \geq N_1 + N_u - 1, \\ &(\mathrm{C2}) \quad N_1 \geq N_B, \\ &(\mathrm{C3}) \quad N_u = N_A + 1. \end{aligned} \tag{44}$$

■

REMARK 9 *Having in mind both the design parsimony (with respect to N_2 and N_u) and the maintenance of the 'essential information' (N_1), two parsimonious (P and S) ways of selecting the observation and control horizons based on the degrees of the plant transfer function N_A and N_B, can be proposed as:*

$$\begin{array}{llcll} (\mathrm{P1}) & N_1 = N_B & & (\mathrm{S1}) & N_1 = N_B + 1 \\ (\mathrm{P2}) & N_2 = N_A + N_B & \textit{and} & (\mathrm{S2}) & N_2 = N_A + N_B + 1 \\ (\mathrm{P3}) & N_u = N_A + 1 & & (\mathrm{S3}) & N_u = N_A + 1 \end{array}$$

In these 'square' tuning settings the effective observation horizon equals the control horizon:

$$N_0 = N_2 - N_1 + 1 = N_A + 1 = N_u \qquad \blacksquare$$

The above presented prime triple-condition (Ci′) and the suboptimal tuning set (S) are a basis for derivation of a numerical algorithm (CD-HAG) presented in Section 4.

3.2. Reducible Polynomials $\widehat{A}(q^{-1})$ & $B(q^{-1})$

Let us presume that

$$\widehat{A}(q^{-1}) = \widehat{A}_0(q^{-1})\Lambda(q^{-1}) \tag{45}$$

$$B(q^{-1}) = B_0(q^{-1})\Lambda(q^{-1}) \tag{46}$$

where $\widehat{A}(q^{-1})$ and $B_0(q^{-1})$ are relatively prime, and $\Lambda(q^{-1})$ denotes the greatest common factor of $\widehat{A}(q^{-1})$ and $B(q^{-1})$ with $\deg \Lambda(q^{-1}) = N_\Lambda > 0$, which will be referred to as the cancellation order. Clearly, the Markov parameters of $\bar{B}(q^{-1})/\widehat{A}(q^{-1})$ are identical to the Markov parameters of $\bar{B}_0(q^{-1})/\widehat{A}(q^{-1})$, where $\bar{B}_0(q^{-1}) = qB_0(q^{-1})$ (see eqn. 9). Therefore it follows from (25)–(26) that the closed-loop characteristic polynomial $D(q^{-1})$ has now the form

$$D(q^{-1}) = C(q^{-1})\Lambda(q^{-1})(\widehat{A}_0(q^{-1}) - q^{-1}\widetilde{A}_0(q^{-1}) + g^* B_0(q^{-1})) \tag{47}$$

where $\widetilde{A}_0(q^{-1}) = \sum_{i=1}^{N_2-N_1+1} k_i(\widehat{A}_0(q^{-1})H_{N_1+i-1}(q^{-1}) - \bar{B}_0(q^{-1}))q^{N_1+i-1}$. By reconsidering the proofs of Theorem 1 and Lemma 1 one can easily find that the following general lemma and theorem hold.

LEMMA 2 *If $\widehat{A}(q^{-1})$ and $B(q^{-1})$ have a common factor of degree N_Λ and one of the two following triple conditions*

(C1′)	$N_2 \geq N_1 + N_A - N_\Lambda$		(C1″)	$N_2 \geq N_1 + N_u - 1$
(C2′)	$N_1 > N_B - N_\Lambda$	*or*	(C2″)	$N_1 = N_B - N_\Lambda$
(C3′)	$N_u = N_A - N_\Lambda + 1$		(C3″)	$N_u \geq N_A - N_\Lambda + 1$

is satisfied, then $\boldsymbol{H}$ is of full column rank. $\blacksquare$

THEOREM 3 *If $\widehat{A}(q^{-1})$ and $B(q^{-1})$ have a common factor of degree N_Λ and one of the two following triple conditions*

(C1′)	$N_2 \geq N_1 + N_u - 1$		(C1″)	$N_2 \geq N_1 + N_u - 1$
(C2′)	$N_1 > N_B - N_\Lambda$	*or*	(C2″)	$N_1 = N_B - N_\Lambda$
(C3′)	$N_u = N_A - N_\Lambda + 1$		(C3″)	$N_u \geq N_A - N_\Lambda + 1$

is satisfied, then $\boldsymbol{H}$ has full column rank ($\operatorname{rank} \boldsymbol{H} = N_u$)*), the solution (11) exists, and, for $\lambda = 0$, the closed-loop characteristic polynomial is*

$$D(q^{-1}) = C(q^{-1})\Lambda(q^{-1})\widetilde{D}_\Lambda(q^{-1}) \tag{48}$$

$$\widetilde{D}_\Lambda(q^{-1}) = 1 + g^* B_0(q^{-1}). \quad \blacksquare \tag{49}$$

REMARK 10 *Note that this time with $r_i = 1$, $i = N_1, \ldots, N_2$, the characteristic polynomial $D(q^{-1})$ set in (25) is only partly determined by the observer polynomial $C(q^{-1})$*

$$D(q^{-1}) = C(q^{-1})\Lambda(q^{-1}). \tag{50}$$

This leads to closed-loop systems stable for all stabilisable systems of (1). On the other hand, for another choice of $\mathbf{r}$ from (19), in fundamental stability considerations, the zeros of the factor $\widetilde{D}_\Lambda(q^{-1})$ have to be examined. To this end the idea of root loci can be applied (Kowalczuk and Suchomski, 1995, Kowalczuk, et al., 1996) to show that for sufficiently small absolute values of g^* the polynomial $\widetilde{D}_\Lambda(q^{-1})$ will always be stable.

Note that the continuous-time approach (Kowalczuk, et al., 1996, Kowalczuk and Suchomski, 1999a,b) offers new alternatives in this respect. ■

REMARK 11 *The conditions of Theorem 3 are analogous to those of Theorem 2, provided that the degrees of the numerator (N_B) and denominator (N_A) of the plant transfer function are reduced by the cancellation order N_Λ. In other words, it is the 'effective degrees' (i.e. the parameters of a minimal realisation of the plant model) that should be appropriately utilised. Note that the value of the cancellation order (N_Λ) is used explicitly in Theorem 3. This is especially disadvantageous in the case of the condition (C2'), which imposes the necessity of precise knowledge of the effective degree of the numerator*

$$N_B^0 = N_B - N_A$$ ■

REMARK 12 *Since the condition (C3') has been chosen (Kowalczuk and Suchomski, 1995) to detect the upper bound on N_u, the first set of conditions (Ci') seems to be a favourable basis for choosing the design parameters. Consequently, with reference to Remark 9, although resulting in the same observation horizon N_0 and the same order of the controller, it is the suboptimal tuning procedure (S) that should be preferred.* ■

4. Numerical algorithms

There is a practical problem in the determination of the cancellation order N_Λ. Note that the cancellation can take place in the controlled plant or be induced by identification of an overparameterised plant model. Since N_Λ diminishes the bound on N_u, the cancellation order can be evaluated – as it has been proposed in Kowalczuk and Suchomski (1995), Kowalczuk, et al. (1996) — by means of detection of the upper bound of N_u that guarantees nonsingularity of $\boldsymbol{H}^T\boldsymbol{H}$.

Taking into account solvability of the design problem for $\lambda = 0$, as it results from Lemma 2 and its first set of conditions (Ci'), where $N_2 \geq N_1 + N_A \geq N_1 + N_A - N_\Lambda$ and $N_1 > N_B \geq N_B - N_\Lambda$ (see also Remark 11), there always exists an upper bound on N_u:

$$N_u \leq N_u^{\max} = N_A - N_\Lambda + 1$$

that establishes the maximum value of the GPC controller design parameter N_u necessary to assure problem solvability. As the cancellation order suitably diminishes the bound on N_u, with an arbitrarily assumed value of N_A the cancellation order can simply be estimated by detecting the maximum value of $N_u^{\max}$:

$$N_\Lambda = N_A - N_u^{\max} + 1 \tag{51}$$

(see also condition C3′ of Theorem 3). The significance of this procedure lies in the fact that the detected parameter $N_u^{\max}$ is an equivalent of the effective system order $N_A^0 = N_A - N_\Lambda$, namely:

$$N_A^0 = N_u^{\max} - 1 \tag{52}$$

Consequently, both of the effective degrees can be found from

$$N_A^0 = N_A - N_\Lambda \tag{53}$$

$$N_B^0 = N_B - N_\Lambda. \tag{54}$$

The process of estimating the bound $N_u^{\max}$ consists in detection of the degeneracy of the 'Markov' matrix (12). Let the matrix $\boldsymbol{H}_{i+1} \in \Re^{N_0 \times (i+1)}$, $i = 1, 2, \ldots$, be partitioned as follows

$$\boldsymbol{H}_{i+1} = [\boldsymbol{H}_i \ \vdots \ \boldsymbol{h}_{i+1}] \tag{55}$$

where

$$\boldsymbol{H}_i = \begin{bmatrix} h_{N_1-1} & h_{N_1-2} & \cdots & h_{N_1-i} \\ h_{N_1} & h_{N_1-1} & \cdots & h_{N_1-i+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_2-1} & h_{N_2-2} & \cdots & h_{N_2-i} \end{bmatrix}, \quad \boldsymbol{H}_i \in \Re^{N_0 \times i},$$

$$\boldsymbol{h}_i = \begin{bmatrix} h_{N_1-i} \\ h_{N_1-i+1} \\ \vdots \\ h_{N_2-i} \end{bmatrix}, \quad \boldsymbol{h}_i \in \Re^{N_0} \text{ and } h_k = 0 \text{ for } k < 0.$$

Consider first the projection $\boldsymbol{I} - \boldsymbol{H}_i \boldsymbol{H}_i^+$ on $\mathcal{R}(\boldsymbol{H}_i)^\perp$, where $\boldsymbol{H}_i^+ \in \Re^{i \times N_0}$ denotes the Moore–Penrose pseudo inverse of $\boldsymbol{H}_i$, and $\mathcal{R}(\boldsymbol{H}_i)^\perp$ stands for the orthogonal complement of the range of $\boldsymbol{H}_i$. Let us assume that, for some i, $\boldsymbol{h}_{i+1}$ belongs to the null space of $\boldsymbol{I} - \boldsymbol{H}_i \boldsymbol{H}_i^+$:

$$\boldsymbol{h}_{i+1} \in \mathcal{N}(\boldsymbol{I} - \boldsymbol{H}_i \boldsymbol{H}_i^+).$$

Hence we can conclude that $\boldsymbol{H}_i \boldsymbol{H}_i^+ \boldsymbol{h}_{i+1} = \boldsymbol{h}_{i+1}$ and $\boldsymbol{h}_{i+1} \in \mathcal{R}(\boldsymbol{H}_i)$. Ultimately

$$\operatorname{rank} [\ \boldsymbol{H}_i \ \vdots \ \boldsymbol{h}_{i+1} \ ] = \operatorname{rank} \boldsymbol{H}_i. \tag{56}$$

From the common rule of computing the pseudo inverse of partitioned matrices (Boullion and Odell, 1971) it results that

$$[\ \boldsymbol{H}_i \ \mid \ \boldsymbol{h}_{i+1}\ ]^{+} = \left[\frac{\boldsymbol{H}_i^{+} - \boldsymbol{H}_i^{+}\boldsymbol{h}_{i+1}\boldsymbol{n}_{i+1}^{+}}{\boldsymbol{n}_{i+1}^{+}}\right] \tag{57}$$

where $\boldsymbol{n}_{i+1} \in \Re^{N_0}$ is defined as

$$\boldsymbol{n}_{i+1} = (\boldsymbol{I} - \boldsymbol{H}_i\boldsymbol{H}_i^{+})\boldsymbol{h}_{i+1}.$$

Assuming that $\boldsymbol{n}_{i+1} \neq \boldsymbol{0}$, and $\boldsymbol{P}_i = \boldsymbol{I} - \boldsymbol{H}_i\boldsymbol{H}_i^{+}$, such that $\boldsymbol{P}_i \in \Re^{N_0 \times N_0}$, one obtains the following recursive solution

$$\boldsymbol{P}_{i+1} = \boldsymbol{P}_i - \frac{\boldsymbol{n}_{i+1}\boldsymbol{n}_{i+1}^{\mathrm{T}}}{\boldsymbol{n}_{i+1}^{\mathrm{T}}\boldsymbol{n}_{i+1}} \tag{58}$$

On the basis of the above derivation the applicable iterative procedure CD-HAG – for Concurrent Determination of the control Horizon N_u And Gain $\boldsymbol{K}$ (11)-can be stated as follows:

PROCEDURE CD-HAG

(*) *Initialisation*: $i = 0$, $\boldsymbol{h}_1 = [\ h_{N_1-1} \ \ \ldots \ \ h_{N_2-1}\ ]^{\mathrm{T}}$

$$\boldsymbol{P}_0 = \boldsymbol{I}$$
$$\boldsymbol{n}_1 = \boldsymbol{P}_0\boldsymbol{h}_1$$
$$\boldsymbol{n}_1^{+} = \|\boldsymbol{n}_1\|_2^{-2}\boldsymbol{n}_1^{T}$$
$$\boldsymbol{H}_1^{+} = \boldsymbol{n}_1^{+}$$
$$\boldsymbol{P}_1 = \boldsymbol{P}_0 - \boldsymbol{n}_1\boldsymbol{n}_1^{+}$$

(*) *Iteration*: $i \leftarrow i + 1$

$$\boldsymbol{h}_i = [\ h_{N_{1-i}} \ \ \ldots \ \ h_{N_2-i}\ ]^{\mathrm{T}}$$
$$\boldsymbol{n}_i = \boldsymbol{P}_{i-1}\boldsymbol{h}_i$$
$$\|\boldsymbol{n}_i\|_2^2 = \boldsymbol{n}_i^{T}\boldsymbol{n}_i \quad (\text{if } \|\boldsymbol{n}_i\|_2^2 < \varepsilon \text{ then go to Termination}),$$
$$\boldsymbol{n}_i^{+} = \|\boldsymbol{n}_i\|_2^{-2}\boldsymbol{n}_i^{T}$$
$$\boldsymbol{p}_i = \boldsymbol{H}_{i-1}^{+}\boldsymbol{h}_i$$
$$\boldsymbol{H}_i^{+} = \left[\frac{\boldsymbol{H}_{i-1}^{+} - \boldsymbol{p}_i\boldsymbol{n}_i^{+}}{\boldsymbol{n}_i^{+}}\right]$$
$$\boldsymbol{P}_i = \boldsymbol{P}_{i-1} - \boldsymbol{n}_i\boldsymbol{n}_i^{+}$$

(*) *Termination*: $N_u = N_u^{\max} = \operatorname{rank}\boldsymbol{H}_i = i$

$$\boldsymbol{K} = \boldsymbol{H}_i^{+} = (\boldsymbol{H}_i^{T}\boldsymbol{H}_i)^{-1}\boldsymbol{H}_i^{T}.$$

where $\boldsymbol{H}_{i+1}$, $i = 1, \ldots,$ *denotes a matrix of the following structure*

$$\boldsymbol{H}_{i+1} = [\boldsymbol{H}_i \ \vdots \ \boldsymbol{h}_{i+1}]$$

and

$$\boldsymbol{H}_i = \begin{bmatrix} h_{N_1-1} & h_{N_1-2} & \cdots & h_{N_1-i} \\ h_{N_1} & h_{N_1-1} & \cdots & h_{N_1-i+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_2-1} & h_{N_2-2} & \cdots & h_{N_2-i} \end{bmatrix},$$

$$\boldsymbol{h}_{i+1} = \begin{bmatrix} h_{N_1-i+1} \\ h_{N_1-i+2} \\ \vdots \\ h_{N_2-i+1} \end{bmatrix}, \ h_k = 0 \text{ for } k < 0.$$

REMARK 13 *It remains to emphasise that with* $\lambda = 0$ *for the terminal* $N_u = \operatorname{rank} \boldsymbol{H}_i = i$ *one obtains the required solution* $\boldsymbol{K}$ *of (11). In a practical case of detection of linear dependency, the terminating condition* $\|\boldsymbol{n}_i\|_2^2 < \varepsilon$ *is applied, where* ε *is a very small computer-dependent real value.* ■

REMARK 14 *It is clear from Theorems 2 and 3 that the second set of conditions (Ci') does not lead to analogous results, because the parameter* N_u *can have an arbitrarily large value.* ■

Assume that r is the first co-ordinate of the vector $\boldsymbol{r}$ that is slightly smaller than one ($r = r_{N_1} \cong 1$), while the other co-ordinates of $\boldsymbol{r}$ are of unit value $r_{N_1+1} = \cdots = r_{N_2} = 1$. A complete AF-GPC design procedure based on the recommended implementation of the anticipation filter with only one parameter (r) submitted to tuning is presented below.

PROCEDURE AF-GPC

1. *Preliminary estimation of the bottom of the observation horizon:* $N_1 = N_B + 1$.
2. *Preliminary estimation of the top of the observation horizon:* $N_2 \geq N_A + N_B + 1$.
3. *Computation of the maximum value of* $N_u = N_u^{\max}$ *with the use of the CD-HAG procedure.*
4. *Calculation of the cancellation order:* $N_\Lambda = N_A - N_u + 1$.
5. *Calculation of the observation horizon's bottom:* $N_1 = N_B - N_\Lambda + 1$.
6. *Calculation of the observation horizon's top:* $N_2 \geq N_1 + N_u - 1$.
7. *Setting a suitable value of the anticipation filter parameter:* $r \in (0.8, 0.99)$.
8. *Computation of the gain matrix* $\boldsymbol{K}$ *and picking its first row* $\boldsymbol{k}^T$.
9. *Solution of the Diophantine equations for the design polynomials* $E_i(q^{-1})$ *and* $G_i(q^{-1})$, *with* $i = N_1, \ldots, N_2$.
10. *Computation of the gain coefficient* g *and the controller polynomials* $L(q^{-1})$ *and* $M(q^{-1})$.

5. Simulation examples

5.1. Simulation settings

The basic form of the plant under examination is given by the following minimal model with relatively prime polynomials $\widehat{A}(q^{-1})$ and $B(q^{-1})$

$$\begin{aligned} A(q^{-1}) &= (1-0.67032q^{-1})(1-0.76593q^{-1})(1-0.81873q^{-1}) \\ B(q^{-1}) &= 0.0028689(1+0.21523q^{-1})(1+3.01224q^{-1})q^{-1} \\ C(q^{-1}) &= (1+0.9q^{-1})[(1+0.63639q^{-1})^2+(0.63639q^{-1})^2]. \end{aligned} \tag{59}$$

The following integral index describes the performance of transient processes

$$I = \sum_{t=0}^{\infty} f(t)^2 = \frac{1}{2\pi j}\oint F(z)F(z^{-1})z^{-1}\mathrm{d}z \tag{60}$$

where $\{f(t)\}_0^\infty$ and $F(z)$ denote the time and the z-domain representations of a given process. In the sequel, two particular cost functions I_u and I_e are considered that refer to the control signals $u(t)$ and the control errors $e(t)$, respectively.

5.2. Elementary performance

According to the first parsimonious rule of selecting the observation and the control horizons (P1–P3), the applied design settings are as follows

$$\begin{aligned} &\lambda = 0,\ N_1 = N_B = 3,\ N_2 = N_A + N_B = 6,\ N_u = N_A + 1 = 4 \\ &r = r_{N_1} \in [0.8, 1]. \end{aligned}$$

The resulting plant input and output signals for the unit step reference signal $w(t)$ are demonstrated in Figs. 1a and 1b, respectively. Clearly, the desired effect of reduction of the control effort is obtained at the cost of a slight deterioration of the transient of the controlled process.

Taking into account the control quality indices I_u and I_u shown in Fig. 1c, a compromising effect of the AF filter implemented with the use of the anticipation parameter $r = r_{N_1}$ is transparent. The plots of the indices exhibit a monotonous character. Therefore, a particular choice of r can be imposed by a need of balancing both of the cost functions.

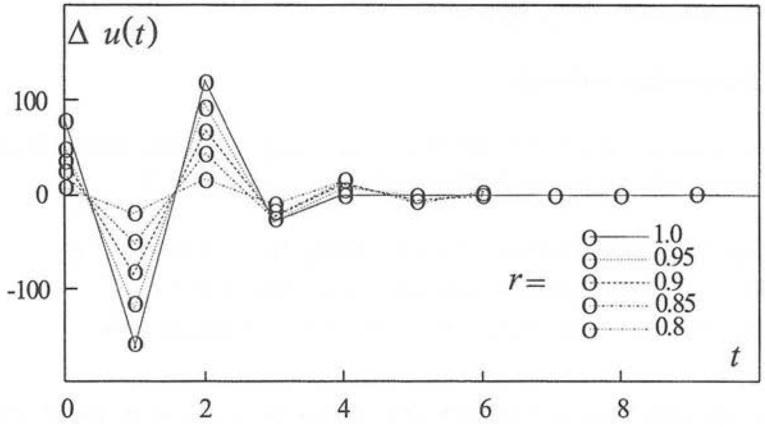

Figure 1a. The control signal for different anticipation (r)

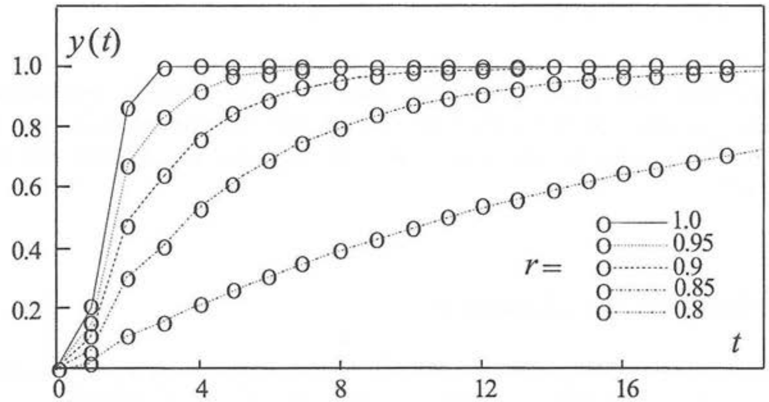

Figure 1b. The effect of anticipation (r) on the system step response

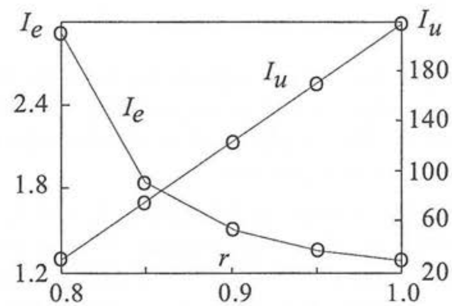

Figure 1c. The control effort (I_u) and error (I_e) indices versus parameter r

Thus, due to the AF mechanism the two fundamental factors of control quality (i.e. control error and control effort) can be taken into account in a rational manner. Note that a balanced relation between I_u and I_e can be obtained in spite of $\lambda = 0$. This means that not only the existence of the solution (11) can be ensured but also the design procedure can be considerably simplified. On the other hand, it should be emphasised that for $\lambda = 0$ there are no apparent conditions for solvability of the design problem and, what is more, frequently with a 'natural' discrepancy between the magnitudes of the control and controlled signals, the λ-tuning can bring about practical difficulties.

5.3. Determination of the observation and control horizons for minimal models

Assume that the controlled plant is described by its minimal model (47) with relatively prime polynomials $\widehat{A}(q^{-1})$ and $B(q^{-1})$, and that the GPC design is performed based on the following rule

$$\lambda = 0, \quad N_1 = N_B = 3, \quad N_2 \geq N_A + N_B,$$
$$N_u = N_0 = N_2 - N_1 + 1, \quad r = 0.9,$$

which results from a necessary modification of the parsimonious rule of the GPC tuning P1–P3.

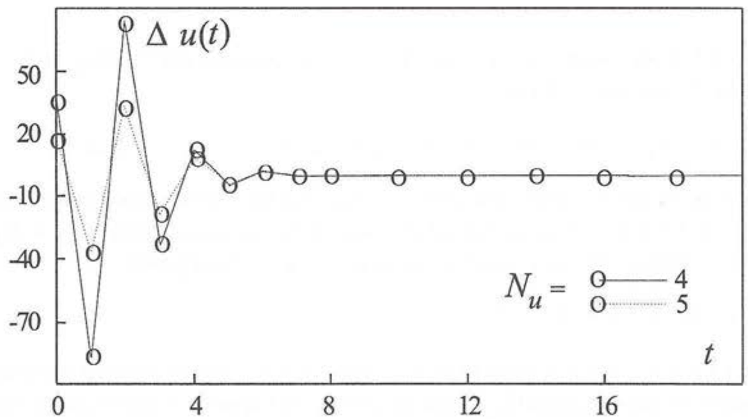

Figure 2a. The control signal for two control horizons (N_u)

As it results from Remark 6, the control horizon N_u should not be greater than the effective observation horizon N_0. On the other hand, if the top instant of the observation horizon N_2 is increased (for example in order to have a more thorough description of the plant) an apparently obvious advice "let N_u follow

N_2" turns out to be disadvantageous. Note that by following the above mentioned advice one obtains a less scrupulous implementation of the computed control sequence $\Delta \boldsymbol{u}(t)$ connected with a decreasing value of the relative range of realisation (RRR) of $\Delta \boldsymbol{u}(t)$. These effects can be seen in Figs. 2a and 2b that show the plots of the control and controlled signals, respectively, corresponding to the unit step change in the setpoint for two values of N_u.

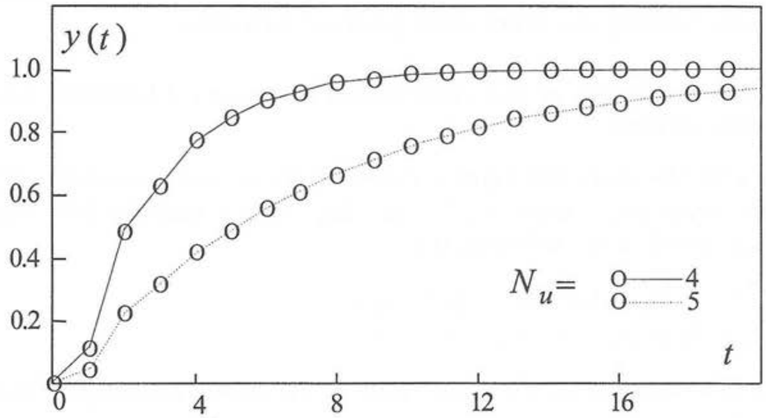

Figure 2b. The influence of the control horizon (N_u) on the system response

The solid lines refer directly to the first parsimonious tuning rule, which leads to the following settings

$$\lambda = 0, \quad N_1 = N_B = 3, \quad N_2 = N_A + N_B = 6, \quad N_u = N_0 = 4.$$

The dotted lines denote the case with the top instant of the observation horizon N_2 chosen to be greater than the above specified minimal value $N_A + N_B$ and the control horizon N_u increased according to the prescription

$$N_u = N_0 = N_2 - N_1 + 1.$$

It is clear that decreasing of the relative range of the control sequence realisation (RRR) leads to an undesirable degradation in the speed of the control process.

In contrast to the above effect it can be shown that by increasing the top instant of the observation horizon N_2 with the control horizon N_u kept constant, the control process can be speeded up. As above, the control and the controlled signals are taken into consideration in a similar experiment with the unit step change of the setpoint. The GPC controller is designed with the following settings

$$\lambda = 0, \quad N_1 = N_B = 3, \quad N_2 \geq N_A + N_B, \quad N_u = N_A + 1 = 4, \quad r = 0.9.$$

According to the P1 and P3 principles both the bottom instant of the observation horizon N_1 and the control horizon length N_u are chosen at their lower bounds. The top instant of the observation horizon N_2 is increased from its lower bound $N_A + N_B$ upward. The simulation results are illustrated in Fig. 3, where Figs. 3a and 3b show the control and output signals obtained for different values of N_2, and the control quality indices and I_u are presented in Fig. 3c.

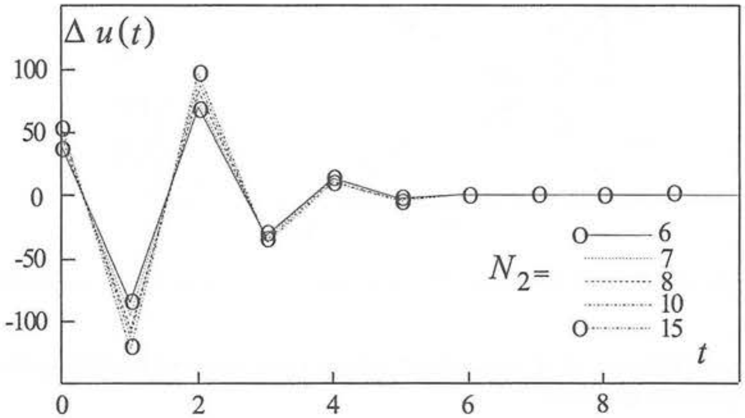

Figure 3a. The control signal for different tops (N_2) of the observation horizon

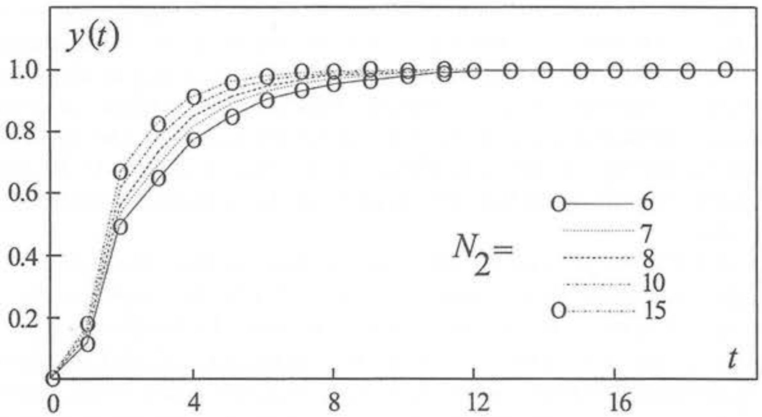

Figure 3b. The control of the observation horizon's top N_2 over the system responses

As it results from these figures, by increasing N_2 a clear acceleration of the control process can be obtained – this can be done, however, along with a slight growth of the control effort. This effect can be interpreted as a result of taking into account more and more exact Markovian representations of the plant while having the RRR ratio frozen at a possibly lowest level. Moreover, it can be notified that the rate of the process acceleration is a monotonically increasing function of the effective observation horizon N_0.

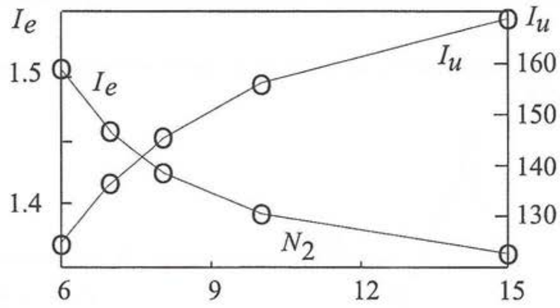

Figure 3c. The control effort (I_u) and error (I_u) indices versus parameter

5.4. Determination of the observation and control horizons for non-minimal models

Consider now the GPC tuning problem for the plant models having the polynomials $\widehat{A}(q^{-1})$ and $B(q^{-1})$ not relatively prime. Note that common factors for the numerator and denominator polynomials of the plant transfer function may arise in the case of identification of plants with variable model orders (and/or delays) by a parameter estimation procedure based on an overparameterised model. The Markovian characterisation of a plant is invariant under any perfect pole-zero cancellation in the transfer function of the plant. However, the GPC design (performed for an overparameterised model of the plant without the exact knowledge of the cancellation order N_Λ) is bound to be based on the assumed (though overestimated) degrees of the numerator and denominator polynomials.

Consider the control system with the anticipated filter characterised by $r = 0.99$. The remaining GPC parameters are established based on the second parsimonious tuning rule (S1–S3). It means that the bottom instant of the observation horizon N_1 is equal to its lower bound ($N_B + 1$) and the top instant of the observation horizon N_2 ($= N_1 + N_u - 1$) is also equal to its lower bound ($N_A + N_B + 1$). The control horizon N_u, in general, appears as a partly free parameter. Nevertheless, irrespective of the assumed estimate of the system orders (the model polynomial degrees) the true upper bound on the control horizon N_u, for which the problem (11) has the required solution, is established by the true plant order $N_u = N_A^0 + 1$ (see C3′ of Lemmas 1 and 2, and eqn. 52).

The step-response performance resulting from overestimated values of the bottom N_1 of the observation horizon is illustrated in Figs. 4a–c, where Figs. 4a and 4b concern the control and output signals of the plant with the solid lines denoting the right choice of $N_A = N_B = 3$, and the plots given in Fig. 4c show the control quality indices I_u and I_e as functions of N_1, i.e. based on the assumed estimates of the degrees of the model polynomials $A(q^{-1})$ and $B(q^{-1})$.

The overestimation of the system order leads to an ascent of the bottom of the observation horizon $N_1 = N_B + 1$ and to an increase the effective observation horizon $N_0 = N_A + 1$. Consequently, such a design procedure causes a degradation in the quality of control, including a reduction of the speed of the controlled process (see Fig. 4b). For comparative purposes, the numerical values of the indices I_u and I_e are also listed in Table 1.

N_1	I_u	I_e
4	185.148	1.3269
5	157.244	1.3894
6	118.386	1.5339
7	62.637	2.0063

Table 1. Numerical values of the cost functions while overestimating N_1.

After an inspection of the above examples one comes to a conclusion that unsatisfactory results of the GPC design caused by inaccurate modelling of the plant can be considerably corrected by appropriately increasing the top instant of the observation horizon N_2. Certainly, if the bottom instant of the observation horizon N_1 and the control horizon N_u are fixed, an optimal shape of the control process can be 'retrieved' by increasing the effective observation horizon N_0.

To illustrate the above proposition let us consider the following design settings

$$\lambda = 0, \quad N_1 = 5, \quad N_2 \geq 9, \quad N_u = 4, \quad r = 0.99.$$

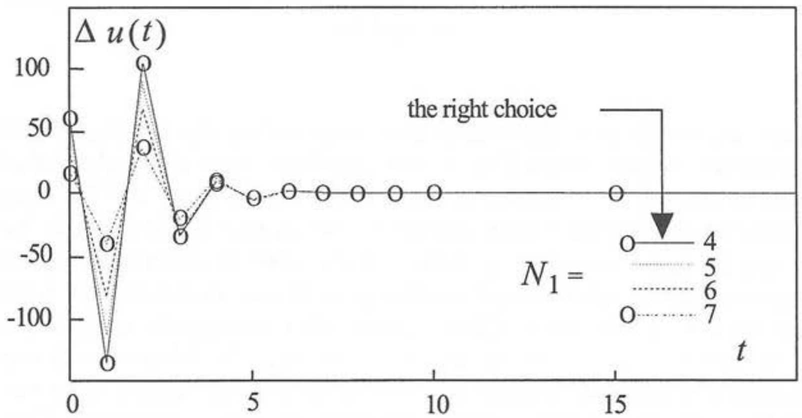

Figure 4a. The control signal for different bottoms (N_1) of the observation horizon in the case of overestimation of the system order

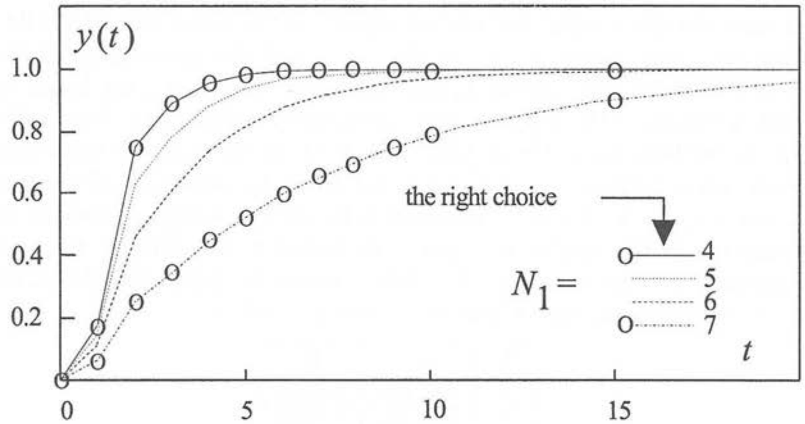

Figure 4b. The effect of the observation horizon's bottom (N_1) on the system reaction invoked by overestimation of the system order

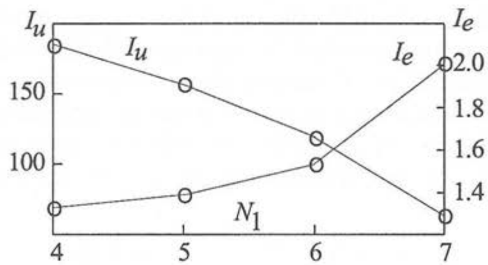

Figure 4c. The effect of order overestimation (N_1) on the control quality items (I_u and I_e)

Figs. 5a and 5b show the control and controlled signals obtained for the unit step setpoint change. From Fig. 5c one can learn more about the influence of the top instant of the observation horizon N_2 on indices I_u and I_e. Numerical values of the considered control quality indices I_u and I_e are listed in Table 2.

It can be easily seen that by increasing the effective observation horizon N_0 an improvement of the transient control process can be obtained at a cost of a certain growth in the control effort. Taking into account the numerical results cited in Tables 1 and 2, it can be concluded that by following the proposed procedure it is possible to make the control error even smaller than the error obtained in the case of the 'correct' modelling previously discussed. This, however, involves more extensive design computations due to the required rise in the top (N_2) of the observation horizon.

N_2	I_u	I_e
9	157.244	1.3894
10	169.518	1.3586
12	182.771	1.3312
15	191.722	1.3157
20	197.645	1.3067

Table 2. Numerical values of cost functions for the tuning of N_2

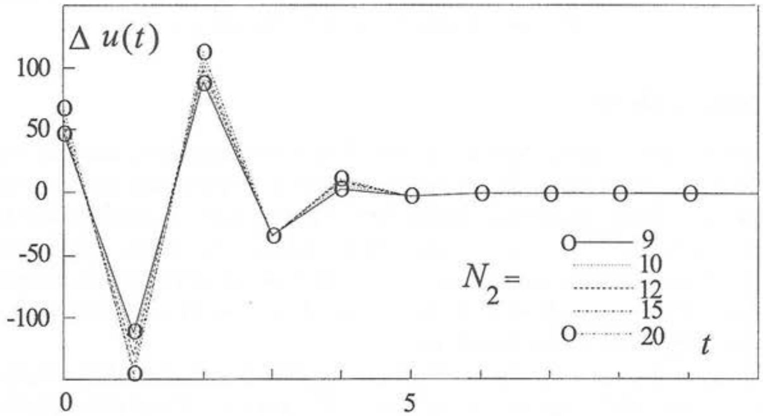

Figure 5a. The control signal for ascended tops (N_2) of the observation horizon in the case of overestimated N_1

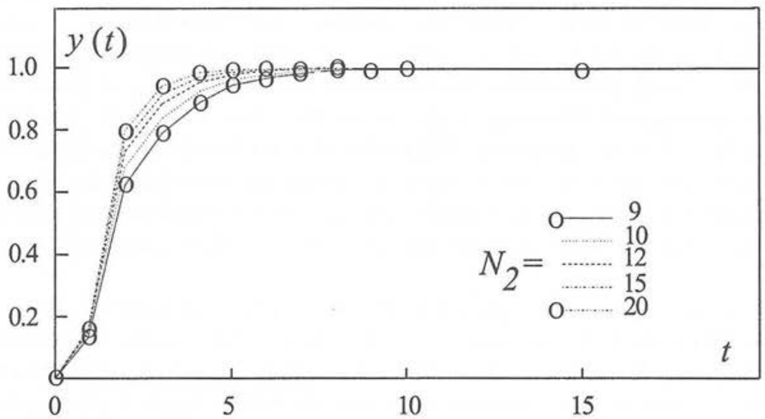

Figure 5b. A remedy for the order overestimation problem: retrieving the control performance with ascended tops (N_2) of the observation horizon

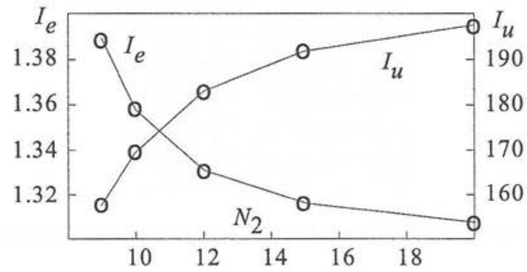

Figure 5c. The case of order overestimation: the effect of ascended observation tops (N_2) on the quality indices (I_u and I_e)

6. Conclusions

It is known that designs based on the dead-beat approach lead to excessive control action, which can in many cases result in a very limited practical range of and some sensitivity problems. It has been shown that the anticipated filtering can have a desired effect on the closed-loop behaviour of the controlled plant (in terms of pole placement, for instance) and that an appropriate design of the anticipated filter can reduce a disagreeable control effort and lead to a certain balance in the control cost functions.

An issue, which is vital for a successful application of the GPC designs, is to provide the user with a set of tuning rules. The proposed iterative algorithm for the simultaneous determination of the control horizon N_u, the matrix controller gain $\boldsymbol{K}$, and the cancellation order, is in line with the latter demand. The developed tools make it possible to use any non-minimal (overestimated) models and, at the same time, to design regulator of a reasonable order.

Simple tuning rules have been analytically obtained due to the assumption that λ is equal to zero. There are, however, no restrictions as to using the non-zero λ in the GPC design solution (11) after the proposed settlement of the design tuning parameters. Nevertheless, one has to keep in mind that the effect of anticipated filtration is similar to the effect of $\lambda \neq 0$. On the other hand, in fact, it is the anticipated filtration that makes the λ-tuning practicable (without it, the value of λ necessary to minimise the cost functions (3) and (17) often appears to be extremely small, and, in effect, unimplementable) and thus allows for a further alleviation of the GPC control effort (originally attributed to λ).

The prospects of the application of the simplest anticipation filter results from mathematical characteristics of the design that consist in reducing the effect of the AF filter to a scalar gain coefficient g. This fact, supported also by experience, leads to the conclusion that it is enough to apply a single parameter r for tuning the AF filter (or, equivalently, the outer loop gain g).

In analytical considerations a constant setpoint is considered, but, principally, this implicitly takes place only in the *anticipation*-time domain. On the other hand, there are no stringent restrictions placed on the input signal, and it is apparent that the discussed design solution can also be used for a tracking problem. It is also worth noticing that the unit step function applied in the reported simulation study is considered in industry as one of the most severe types of excitation signals.

References

BOULLION, T.L. and ODELL, P.L. (1971) *Generalized Inverse Matrices.* New York, Wiley-Interscience, John Wiley & Sons.

CLARKE, D.W., MOHTADI, C. and TUFFS, P. S. (1987) Generalized predictive control — Part I and Part II. *Automatica*, **23**, 2, 137–160.

CLARKE, D.W. and MOHTADI, C. (1987) Properties of generalised predictive control. *Proc. of the Tenth IFAC World Congress*, Munich, Germany, **10**, 63–74.

DEMIRCIOGLU, H. and GAWTHROP, P.J. (1991) Continuous-time generalised predictive control. *Automatica*, **27**, 1, 55–74.

FAVIER, G. (1987) Self-tuning long-range predictive controllers. *Proc. of the Tenth IFAC World Congress*, Munich, Germany, **10**, 80–87.

GOREZ, R., WERTZ, V. and ZHU, K.-Y. (1987) On a generalised predictive control algorithm. *Systems and Control Letters*, **9**, 5, 369–377.

KOWALCZUK, Z. and SUCHOMSKI, P. (1995) Anticipated filtering approach to generalised predictive control. *Proc. of the European Control Conference*, Rome, Italy, **4**, 3591–3596.

KOWALCZUK, Z. and SUCHOMSKI, P. (1996) Discrete-time generalised predictive control with anticipated filtration of control error. *Proc. of the 1996 Triennial World Congress of IFAC*, San Francisco, USA, **K**, 301–306.

KOWALCZUK, Z. and SUCHOMSKI, P. (1999a) Control of delay plants via continuous-time GPC principle. *Control and Cybernetics*, **28**, 2, 291–314.

KOWALCZUK, Z. and SUCHOMSKI, P. (1999b) Continuous-time generalised predictive control of delay systems. *IEE Proceedings — Control Theory and Applications*, **146**, 1, 65–75.

KOWALCZUK, Z., SUCHOMSKI, P. and MARCIŃCZYK, A. (1996) Discrete-time and continuous-time generalised predictive controllers with anticipated filtration: Tuning rules. *Applied Mathematics and Computer Science*, **6**, 4, 707–732.

ZHANG, J. (1996) Property analysis of GPC based on coefficient mapping. *Proc. of the 1996 Triennial World Congress of IFAC*, San Francisco, USA, **C**, 457–462.