15th Triennial World Congress, Barcelona, Spain

IMPROVED PI CONTROL VIA DYNAMIC INVERSION

Aurelio Piazzi * **Antonio Visioli** **

* *Dipartimento di Ingegneria dell'Informazione, University of Parma (Italy), e-mail: aurelio@ce.unipr.it*
** *Dipartimento di Elettronica per l'Automazione, University of Brescia, (Italy), e-mail: visioli@bsing.ing.unibs.it*

Abstract: In this paper we present a new design methodology to achieve high performances in the setpoint following task by means of a standard Proportional-Integral (PI) controller. In particular, a dynamic inversion procedure is adopted to synthesize a suitable command input to the closed-loop control system. In this way, the PI parameters can be selected for a load rejection purpose, as the designed command input is capable to ensure in any case that a desired output transition is accomplished with a low rise time and a small overshoot. A salient feature of the technique is that it does not require a complex identification phase, as a simple first-order plus time delay (FOPTD) model of the process, which can be easily obtained by means of an open-loop step response, suffices. A Padé approximation is then adopted to deal with time delays. Illustrative examples show the effectiveness of the method, which preserves all the basic characteristics of the PI controllers and it is therefore suitable to be applied in an industrial environment.

Keywords: PI control, dynamic inversion, setpoint following.

1. INTRODUCTION

Proportional-Integral (PI) controller are undoubtedly the most adopted controllers in industrial settings, because of their capability to provide satisfactory performances for a wide range of plants, despite their simplicity. In fact, for many processes the derivative term of the controller is not useful and it is often difficult to tune, so that practitioners prefer to avoid its use (Aström and Hägglund, 2000*b*).
Many tuning formulas have been provided in the last sixty years to help the operator to find suitable values of the PI parameters depending on the process dynamics (see (O'Dwyer, 2000) for an excellent collection of them). Nevertheless, because of the clear physical meaning of the parameters, in many cases the values of the controller gains are selected by hand from the operator. As a result, the controller performances heavily depend on the operator skilfullness and it might happen that they are far away from the optimality. Besides, in many situations the operator has to face the dilemma to tune the controller in order to achieve either a good setpoint following or a good load rejection. In general, the PI gains are properly tune for the latter task and the setpoint following performances are somehow recovered by filtering the setpoint signal or by applying a setpoint weighting (Shinskey, 1996). However, it is obvious that in any case a decreasing in the performances in the setpoint following task has to be expected.
In this paper we propose a new design methodology for the attainment of high performances in both the load rejection and the setpoint following task. In particular, the proposed method aims at finding a command input to the closed-loop system (in which the PI controller has been properly tuned to guarantee the load rejection specifications) in order to perform a desired output transition (from a setpoint value to another) with no overshoot and with a defined rise time.
Basically, the technique requires the modelling of the plant with a first-order plus time delay (FOPTD) transfer function (which can be obtained by simply evaluating an open-loop step response) in which the time delay is expressed by a Padé approximation in order to have a rational expression. On this basis, the PI

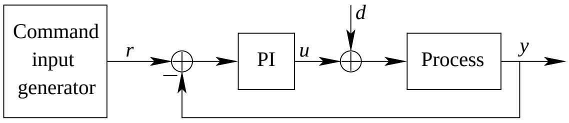

Fig. 1. The overall control scheme.

controller is tuned by any conventional technique that ensures satisfactory performances in the load rejection task (see for example (Zhuang and Atherton, 1993)). Then, a desired output function to be followed in the response of a setpoint change is defined as a "transition" polynomial (Piazzi and Visioli, 2001*b*) and finally the new command input to be applied to the closed-loop system is determined by adopting a stable dynamic inversion technique (Piazzi and Visioli, 2001*a*). Note that this command input replaces the conventional step signal that is very often adopted to attain a new setpoint value.

2. METHODOLOGY

The design approach we propose refers to the control scheme shown in Figure 1, where good performances are required in the rejection of the load disturbance d as well as in the transition of the system output y from one value to another. The method proposed in this paper basically relies in finding the command signal r that provides an efficient system output transition with an already tuned PI controller.

2.1 *Modeling*

The modelling phase is based on a standard open-loop step response evaluation in order to determine a FOPTD model of the plant. Specifically, the well-known area method (Aström and Hägglund, 1995), which is somewhat robust to the measurement noise, can be adopted to determine a transfer function of the process that can be written as:

$$P(s) = \frac{K}{Ts+1}e^{-Ls} \tag{1}$$

To obtain a rational transfer function, it has been chosen to write the exponential term expressing the time delay as a third order Padé approximation, i.e.:

$$e^{-Ls} \cong \frac{1-0.5Ls+L^2s^2/10-L^3s^3/120}{1+0.5Ls+L^2s^2/10+L^3s^3/120}$$

The choice of the order of the Padé approximation has been motivated by the need to have a good approximation in a sufficient range of frequencies. In this way, the approximated process transfer function results to be:

$$\tilde{P}(s) \cong \frac{K}{Ts+1}\frac{1-0.5Ls+L^2s^2/10-L^3s^3/120}{1+0.5Ls+L^2s^2/10+L^3s^3/120}. \tag{2}$$

L/T range	0.1-1.0	1.1-2.0
a_1	1.015	1.065
b_1	-0.957	-0.673
a_2	0.667	0.687
b_2	-0.552	-0.427

Table 1. Tuning rules for the minimization of the ISTE criterion for load rejection.

2.2 *Tuning the PI controller*

In order to apply the dynamic inversion based methodology that will be presented in the next, the PI controller can be tuned according to any of the many methods proposed in the literature or even by hand. However, since the purpose of the dynamic inversion procedure is the attainment of high performances in the setpoint following task, disregarding of the controller gains, it is sensible to select the PI parameters aiming only at obtaining good load rejection performances.
The PI controller transfer function be denoted as follows:

$$C(s) = K_p\left(1+\frac{1}{T_is}\right) \tag{3}$$

where K_p is the proportional gain and T_i is the integral time constant. Among the several tuning formulas devised for the load rejection purpose, we highlight two of them, which will be employed in the illustrative examples (see Section 3). The first is the well known Ziegler-Nichols one ($K_p = \frac{0.9T}{KL}$ and $T_i = 3L$); the other one has been proposed in (Zhuang and Atherton, 1993) and aims at minimizing the ISTE criterion, defined as

$$J = \int_0^\infty te(t)dt$$

where $e(t)$ is the system error (i.e. the difference between the step setpoint value and the current output). To this purpose, analytical tuning rules have been provided. They are:

$$\begin{aligned} K_p &= \frac{a_1}{K}\left(\frac{L}{T}\right)^{b_1} \\ \frac{1}{T_i} &= \frac{a_2}{T}\left(\frac{L}{T}\right)^{b_2} \end{aligned} \tag{4}$$

where, for convenience, the values of a_1, b_1, a_2, b_2 are reported in Table 1.

2.3 *Output function design*

As a desired output function that defines the transition from a setpoint value y_0 to another y_1 (to be performed in the time interval $[0,\tau]$) we choose a "transition" polynomial (Piazzi and Visioli, 2001*b*), i.e a polynomial function that satisfies boundary conditions and that is parameterized by the transition time τ. In the following, without loss of generality we will assume $y_0 = 0$, Formally, define

$$y(t) = c_{2p+1}t^{2p+1} + c_{2p}t^{2p} + \cdots + c_1 t + c_0$$

The polynomial coefficients can be uniquely found by solving the following linear system, in which boundary conditions at the endpoints of interval $[0, \tau]$ are imposed:

$$\begin{cases} y(0) = 0; \quad y(\tau) = y_1 \\ y^{(1)}(0) = 0; \; y^{(1)}(\tau) = 0 \\ \vdots \\ y^{(p)}(0) = 0; \; y^{(p)}(\tau) = 0 \end{cases}$$

The results can be expressed in closed-form as follows ($t \in [0, \tau]$):

$$y(t;\tau) = y_1 \frac{(2p+1)!}{p!\tau^{2p+1}} \sum_{i=0}^{p} \frac{(-1)^{p-i}}{i!(p-i)!(2p-i+1)} \tau^i t^{2p-i+1}. \quad (5)$$

The order of the polynomial can be selected by imposing the order of continuity of the command input that results from the dynamic inversion procedure (Piazzi and Visioli, 2001*b*). Specifically, if the plant is modelled as a FOPTD transfer function (see (1)), its relative degree is equal to one. Taking into account that the relative degree of the PI controller is zero, the relative degree of the overall closed-loop system is one. Thus, a third order polynomial ($p = 1$) suffices if a continuous command input function is required. Outside the interval $[0, \tau]$ the function $y(t;\tau)$ is equal to 0 for $t < 0$ and equal to y_1 for $t > \tau$.

2.4 *Determining the command input via dynamic inversion*

At this point we address the problem of finding the command signal $r(t;\tau)$ that provides the desired output function (5). The closed-loop transfer function be denoted as

$$T(s) := \frac{C(s)\tilde{P}(s)}{1 + C(s)\tilde{P}(s)} = \frac{b(s)}{a(s)}. \quad (6)$$

As $T(s)$ is nonminimum phase, a stable dynamic inversion procedure is necessary, that is a bounded input function has to be found in order to produce the desired output. Denote the set of all cause/effect function pairs $(r(\cdot), y(\cdot))$ associated to $T(s)$ by $\mathcal{B}$. Now, in order to perform the stable inversion, we rewrite the numerator of the transfer function (6) as follows:

$$b(s) = b_-(s)b_+(s)$$

where $b_-(s)$ and $b_+(s)$ denote the polynomials associated to the zeros with negative real part and positive real part respectively (no purely imaginary zeros are present having chosen a PI controller (see (2) and (3)).

Now, consider the inverse system of (6) whose transfer function can be written as:

$$T(s)^{-1} = \gamma_0 + \gamma_1 s + H_0(s)$$

where γ_0 and γ_1 are real constants and $H_0(s)$, a strictly proper rational function, represents the zero dynamics. This can be uniquely decomposed according to

$$H_0(s) = H_0^-(s) + H_0^+(s) = \frac{c(s)}{b_-(s)} + \frac{d(s)}{b_+(s)}$$

where $c(s)$ and $d(s)$ are suitable polynomials. Being $\mathcal{L}$ the Laplace transform operator, define:

$$\eta_0^-(t) := \mathcal{L}^{-1}[H_0^-(s)]$$

$$\eta_0^+(t) := \mathcal{L}^{-1}[H_0^+(s)]$$

$$Y(s;\tau) := \mathcal{L}[y(t;\tau)]$$

The unstable reference function $r_u(t;\tau)$ that causes the desired output function $y(t;\tau)$ can be simply determined as:

$$\begin{aligned} r_u(t;\tau) &= \mathcal{L}^{-1}[T(s)^{-1}Y(s;\tau)] \\ &= \gamma_0 y(t;\tau) + \gamma_1 y^{(1)}(t;\tau) + \int_0^t \eta_0^-(t-v)y(v;\tau)dv \\ &+ \int_0^t \eta_0^+(t-v)y(v;\tau)dv \quad t \in (-\infty, +\infty). \end{aligned} \quad (7)$$

Thus, we have that $(r_u(t;\tau), y(t;\tau)) \in \mathcal{B}$ and note that $r_u(t;\tau) = 0$ if $t \in (-\infty, 0)$ and $r_u(t;\tau)$ is unbounded over $[0, +\infty)$ due to the unstable zero dynamics (associated to $H_0^+(s)$).
The unstable modes associated with $b_+(s)$ be denoted by $m_i(t)$, $i = 1, \ldots, w$. Then, the following lemma results:

Lemma 1. There exists real constants $k_i \in \mathbb{R}$, $i = 0, \ldots, w$ that depend on positive time parameter τ such that, for $t > \tau$

$$\int_0^t \eta_0^+(t-v)y(v;\tau)dv = k_0(\tau) + \sum_{i=1}^{w} k_i(\tau)m_i(t).$$

Proof. Considering that $t > \tau$ we can rewrite the integral of the above Lemma as follows:

$$\begin{aligned} &\int_0^t \eta_0^+(t-v)y(v;\tau)dv = \\ &\int_0^\tau \eta_0^+(t-v)y(v;\tau)dv + \int_\tau^t \eta_0^+(t-v)y_1 dv. \end{aligned} \quad (8)$$

As it is known $\eta_0^+(t)$, the impulse response of the unstable zero dynamics can be expressed as a linear combination of the modes $m_i(t)$:

$$\eta_0^+(t) = \alpha_1 m_1(t) + \cdots + \alpha_w m_w(t) \quad (9)$$

where $\alpha_i \in \mathbb{R}$, $i = 1, \ldots, w$ are appropriate coefficients. Taking into account the analytic expression of the transition polynomial $y(t;\tau)$ it then follows that the integral $\int_0^\tau \eta_0^+(t-v)y(v;\tau)dv$ is a linear combination of the modes $m_i(t)$ and its coefficients depend on τ. On the other hand, examining the integral $\int_\tau^t \eta_0^+(t-v)y_1 dv$ we analogously deduce that it can be expressed as a linear combination of the modes $m_i(t)$ plus a constant addend. Therefore, by virtue of (8) the statement of Lemma 1 follows. □

At this point, taking into account Lemma 1 we can define the following function:

$$r_c(t;\tau) := -\sum_{i=1}^{w} k_i(\tau) m_i(t) \quad t \in (-\infty, +\infty). \tag{10}$$

The following lemma results:

Lemma 2. Being $r_c(t;\tau)$ the function defined in (10), we have

$$(r_c(t;\tau), 0) \in \mathcal{B} \quad \forall \tau \in \mathbb{R}.$$

Proof. By examination of the differential equation associated to the system described by the transfer function $G(s) = (b_-(s)b_+(s))/a(s)$ it follows that the pair $(r_c(t;\tau), 0)$ satisfies this equation over $(-\infty, +\infty)$. □

Finally, we can define the following function, which perform the exact stable inversion:

$$r(t;\tau) = r_u(t;\tau) + r_c(t;\tau) \quad t \in (-\infty, +\infty). \tag{11}$$

The following proposition can therefore be stated.

Proposition 3. The function function $r(t;\tau)$ defined in (11) is bounded over $(-\infty, +\infty)$ and $(r(t;\tau), y(t;\tau)) \in \mathcal{B}$.

Proof. Taking into account Lemma 2, it is evidently $(r(t;\tau), y(t;\tau)) \in \mathcal{B}$ by virtue of linear superposition. On the other hand, $r(t;\tau)$ is, by construction, bounded because of the exact cancellation of all the unstable modes appearing in $r_u(t;\tau)$ (see Lemma 1 and definition (10)). □

Summarizing, the determined function $r(t;\tau)$ exactly solves the stable inversion problem for a family of output functions, which depend on the free transition time τ. This transition time can be arbitrarily selected by the user depending on the required performances and on the saturation level of the actuator.

Actually, from a practical point of view, in order to use the synthesized function (11) it is necessary to truncate it, resulting therefore in an approximate generation of the desired output $y(t;\tau)$. This can be done with arbitrarily precision given any couple of small parameters $\varepsilon_0 > 0$ and $\varepsilon_1 > 0$. Indeed, compute

$$t_0 := \max\{t' \in \mathbb{R} : |r(t;\tau)| \le \varepsilon_0 \;\; \forall t \in (-\infty, t']\}$$

and define

$$t_s := \min\{0, t_0\}.$$

Similarly, compute

$$t_f := \min\{t' \in \mathbb{R} : |r(t;\tau^*) - y_1| \le \varepsilon_1 \;\; \forall t \in [t', \infty)\}$$

Hence, the approximate reference signal to be actually used is

$$r_a(t;\tau^*) := \begin{cases} 0 & \text{for } t < t_s \\ r(t;\tau^*) & \text{for } t_s \le t \le t_f \\ y_1 & \text{for } t > t_f. \end{cases}$$

Note that t_s depends on τ and it might occur that $t_s < 0$, resulting in the so-called "preaction control" (Marro and Piazzi, 1996).

3. ILLUSTRATIVE EXAMPLES

3.1 *FOPTD system*

As a first example we consider a FOPTD system, namely:

$$P(s) = \frac{1}{10s+1} e^{-5s}.$$

The PI controller gains have been selected by applying the Ziegler-Nichols formulas. It results $K_p = 2$ and $T_i = 15.02$. The system output has to perform a transition from 0 to $y_1 = 1$. We fixed the transition time $\tau = 20$. Consequently, the desired output function, according to (5) with $p = 1$ is:

$$y(t) = -\frac{2}{20^3} t^3 + \frac{3}{20^2} t^2.$$

By considering the Padé approximation, so that

$$\tilde{P}(s) = \frac{1}{10s+1} \frac{1 - 2.5s + 25/10s^2 - 125/120s^3}{1 + 2.5s + 25/10s^2 + 125/120s^3}$$

and by applying the dynamic inversion procedure to the resulting closed-loop system (with $\varepsilon_0 = \varepsilon_1 = 10^{-4}$) we obtained the input command function shown in Figure 2 in which it is evident the resulting preaction time $t_s = -10$s (note that, for convenience, the time axis has been properly shifted in order to have $t_s = 0$). The obtained system output is reported in Figure 3 with the corresponding control signal.

It appears that the use of the designed command input function allows to perform a fast transition with a very small overshoot. The improvement achieved with respect to the standard technique in which a step setpoint signal is applied to the closed-loop system is evident by evaluating the system output and the control signal in this case, which are depicted in Figure 3 as well. It can be easily seen that the settling time, as well as the control effort, is much higher than with the dynamic inversion based methodology.

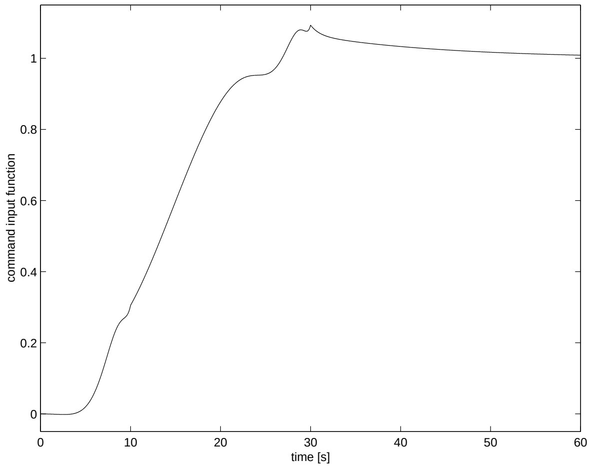

Fig. 2. The optimal reference command input function for the FOPTD system.

3.2 *High order system*

In order to evaluate the effectiveness of the new approach in the presence of a rough approximation of the system dynamics, we consider the high-order process (Aström and Hägglund, 2000*a*)

$$P(s) = \frac{1}{(s+1)^8}.$$

First, we modelled the process with a FOPTD transfer function, by means of the area method, yielding $K = 1$, $T = 3.03$ and $L = 4.96$, i.e.:

$$\tilde{P}(s) = \frac{1}{3.03s+1}\frac{1-2.48s+2.46s^2-1.02s^3}{1+2.48s+2.46s^2+1.02s^3}$$

Then, we tuned the PI controller by means of the Ziegler-Nichols formulas (it results $K_p = 0.61$ and $T_i = 14.90$) and we applied the dynamic inversion based methodology. The time to perform a transition from 0 to 1 has been selected as the settling time (at 2%) of the open-loop system, which is equal to 14.82s. The determined command input function (with $\varepsilon_0 = \varepsilon_1 = 10^{-3}$) is shown in Figure 4. The preaction time is $t_s = -8.9$s. The corresponding system output and control signal are reported in Figure 5.
Analogously, we also applied the tuning method based on the minimization of the ISTE criterion (it results $K_p = 0.76$ and $T_i = 14.81$). With the same value of τ as before we determined the command input function shown in Figure 6 (the preaction time is $t_s - 9$s). System output and control signal are plotted in Figure 7. To better evaluate the significance of the results, in Figure 8 it is plotted the system output with both the PI tuning methods when a step signal is applied to the setpoint at time $t = 0$ and to the load at time $t =$ 200s. It appears that the new approach outperforms the classic one, disregarding the fact that different tuning methods (which produces very different performances in the setpoint following and in the load rejection task) are adopted.

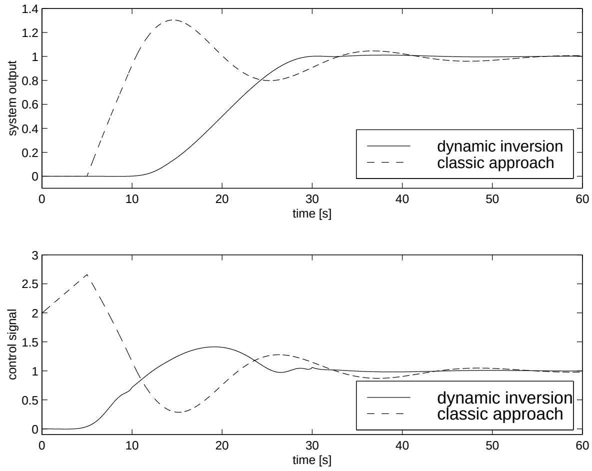

Fig. 3. The system output and the control signal for the FOPTD system (dynamic inversion and classic approach).

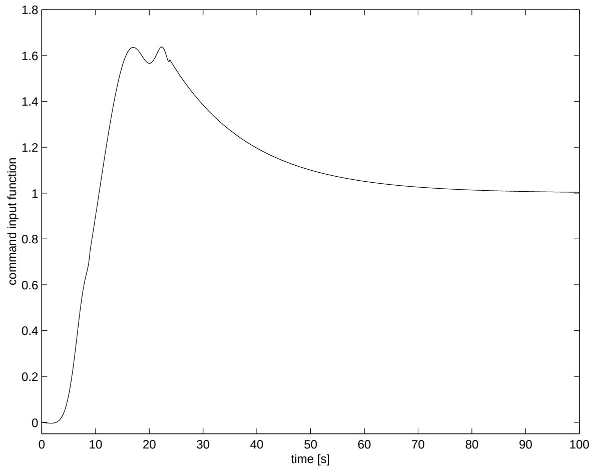

Fig. 4. The optimal command input function for the high order system (Z-N tuning).

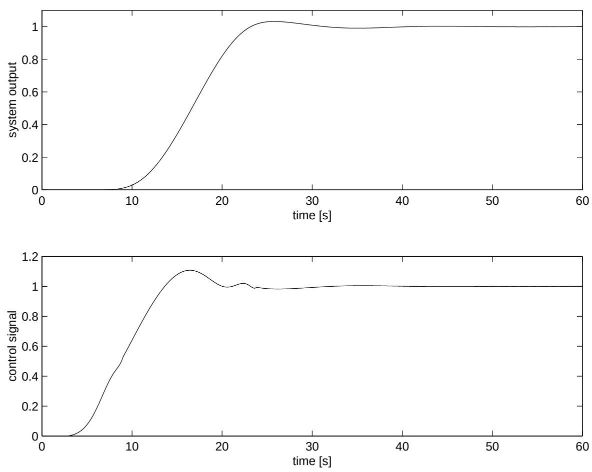

Fig. 5. The system output and the control signal for the high order system (Z-N tuning).

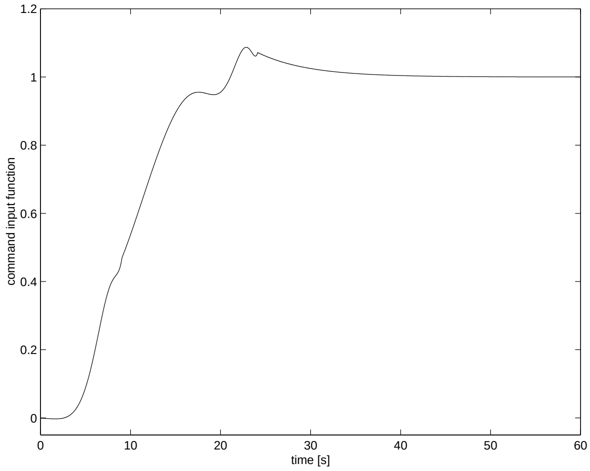

Fig. 6. The optimal command input function for the high order system (ISTE tuning).

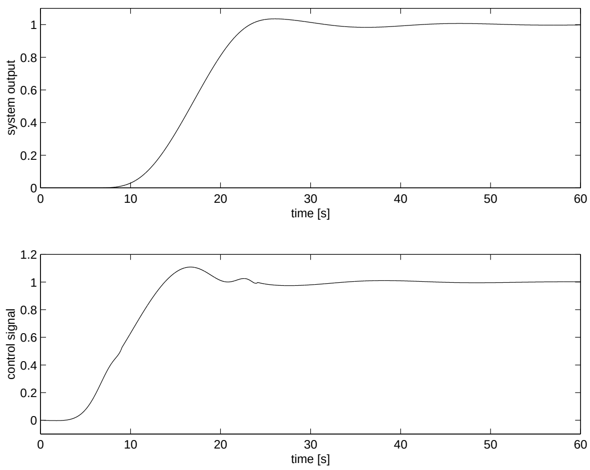

Fig. 7. The system output and the control signal for the high order system (ISTE tuning).

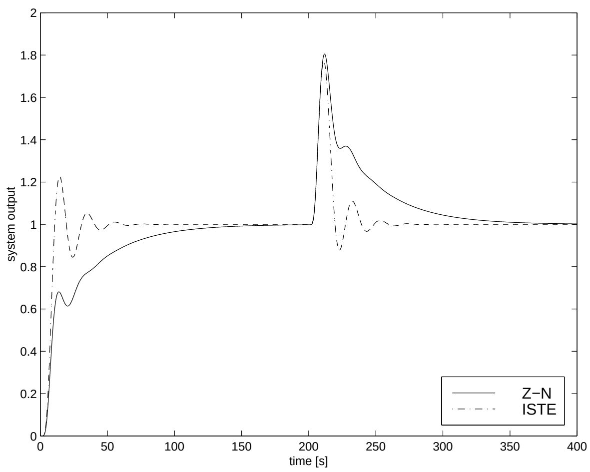

Fig. 8. The system output for the high-order system (classic approach).

4. CONCLUSIONS

In this paper we have presented a new approach, based on dynamic inversion, for the improvement of the setpoint following task of a PI controller. It has been shown that the control schemes in which the proposed technique is adopted outperform the classic ones, disregarding the tuning method employed to set the PI parameters and the fact that the process is simply modelled with a FOPTD transfer function. In this way, both the load rejection and the setpoint following requirements can be satisfied. As the dynamic inversion procedure is applied to the standard closed-loop, scheme the know-how in the general use of the PI controller is fully retained and therefore the approach appears to be suitable to be used in industrial environment, where the operator can set the free design parameter τ to deal with actuator constraints.

Future work will be devoted to extend the technique to PID controller and to the adoption of more complex identification method to improve the performances.

5. REFERENCES

Aström, K. and T. Hägglund (1995). *PID controllers: theory, design and tuning*. ISA Press.

Aström, K. and T. Hägglund (2000*a*). Benchmark systems for PID control. In: *Preprints IFAC Workshop on Digital Control PID‘00*. Terrassa (E). pp. 181–182.

Aström, K. and T. Hägglund (2000*b*). The future of PID control. In: *Preprints IFAC Workshop on Digital Control PID‘00*. Terrassa (E). pp. 19–30.

Marro, G. and A. Piazzi (1996). A geometric approach to multivariable perfect tracking. In: *Proc. of the 13th IFAC World Congress*. San Francisco, CA. pp. 241–246.

O’Dwyer, A. (2000). PI and PID controller tuning rules for time delay processes: a summary. Technical Report 2000-01. School of Control Systems and Electrical Engineering, Dublin Institute of Technology, Dublin (Ireland).

Piazzi, A. and A. Visioli (2001*a*). Optimal inversion-based control for the set-point regulation of nonminimum-phase uncertain scalar systems. *IEEE Trans. on Automatic Control* **46**(10), 1654–1659.

Piazzi, A. and A. Visioli (2001*b*). Optimal noncausal set-point regulation of scalar systems. *Automatica* **37**(1), 121–127.

Shinskey, F. G. (1996). *Process Control Systems - Application, Design, and Tuning*. McGraw-Hill. USA.

Zhuang, M. and D. P. Atherton (1993). Automatic tuning of optimum PID controllers. *IEE Proceedings - Control Theory and Applications* **140**(3), 216–224.