PI(D) tuning with contextual model identification

Alberto Leva, Sara Negro, Alessandro Vittorio Papadopoulos

Abstract— **This work proposes a methodology to deal with model-based tuning of PI(D) regulators, and with the previous identification of the required process model, in a contextual way. The methodology can be applied to existing tuning rules, as done herein, leading to better control results, and also to process models capable of forecasting the closed-loop transients in a reliable manner. The same idea can be generalised to virtually any regulator structure, and also be employed to devise new tuning (and contextual identification) rules—a subject left however to future works.**

I. INTRODUCTION

In Model-Based AutoTuning (MBAT for short) of industrial regulators, some process input/output measurements are first used to obtain a "model" of the process itself; subsequently, the obtained model is employed, together with convenient specifications, to compute the parameters of the regulator with some "tuning rules". With respect to the more general domain of autotuning, peculiar to MBAT is the explicit presence of a process model. There is a vast literature on MBAT, impossible to review here. The interested reader can refer to the short survey [1], that also points out the main advantages (and potential pitfalls) of MBAT, to more extensive works such as [2], [3], [4], and for a broad *panorama* to the excellent survey [5].

In general, MBAT methods are based on very simple process models, the structure of which is typically decided *a priori*, based essentially on the structure of the regulator to be tuned, and sometimes (but not so frequently) on some facts relative to the process dynamics that can be guessed based on the measured data.

The reasons for such a *scenario* are numerous, but at the level of this work, can be summarised as follows. First, in the typically available data there is limited information. In real-world cases, models for MBAT have to be identified on-line, frequently on the basis of small sets of noisy data, produced by *stimuli* that must obey to potentially severe process upset constraints, and therefore typically lack excitation. Second, to achieve tuning procedures suitable for automated implementation in industrial products, it is highly desirable that those models allow for *explicit* regulator parametrisation formulæ [6], [7].

Among the consequences of that situation, three are worth noticing. First, with typical MBAT-compatible models, capturing the control-relevant process dynamics may turn out to be quite tricky. One needs a model that is precise near the cutoff frequency, but that frequency is a *result* of the tuning, thus not yet known when the model is identified. Second, the presence of unmodelled dynamics is an ubiquitous problem, to be accounted for somehow when devising tuning rules. Third, the difficulties encountered in setting up a well-posed identification problem motivate the widespread use of *ad hoc* identification methods such as the method of areas, the method of moments, the tangent method, and so forth.

In any case, the results of MBAT end up depending up to a large extent on the particular method used to identify the used model, and this is certainly a relevant drawback. Indeed, *if the scope is limited to the MBAT context*, there is little point both in discussing a tuning procedure without taking into account the possible effects of the identification method [8], [9], and in devising complex identification methods under hypotheses that are almost impossible to check on the basis of field data: in that context, well established results of the identification theory can often prove inadequate to assess the quality of a model, as shown e.g. in [10]. Finally, the mentioned limitations on the applicable process *stimuli* leave limited or no room for experiment design, so that MBAT can take little (if any) profit of the neat results of the "identification for control" research [11], [12].

The main contribution of this work, that is part of a long-term research path initiated by [13], [14] and continued by [15], is the definition of a methodology capable of complementing any existing model-based tuning rule with a *contextual* method to identify the required process model, so as to provide a solution for the problem sketched above. The focus is here restricted to the PI structure and to the use of frequency-domain data (more precisely, measured points of the process Nyquist curve), as briefly discussed in section II, but the idea as expressed in section III and used in section IV is far more general, and will be further exploited in the future: some directions envisaged for that, after the commented examples of section VII, are reported in section IX.

II. BRIEF PRELIMINARY DISCUSSION

Having established that MBAT poses a peculiar type of identification problem, see e.g. the discussion in [15], the question one has to answer *prior* to initiating the process stimulation is "where (in the broadest sense of the term) is the control-relevant information". Such a question has apparently no general answer, but if one tries to bring in and formalise the huge amount autotuning experience available in the literature and in applications, some useful considerations arise. First, seeking information in the time domain (e.g., the observed delay, time constant and gain of

A. Leva is with the Dipartimento di Elettronica e Informazione, Politecnico di Milano, Via Ponzio 34/5, 20133 Milano, Italy, leva@elet.polimi.it.

S. Negro and A.V. Papadopoulos are graduate students at the Dipartimento di Elettronica e Informazione

a step response) invariantly leads to catch the low-frequency process dynamics, and taking that dynamics *a priori* as the control-relevant one may be *very* misleading—a conclusion that the reader can easily draw from the discussion in [9]. On the other hand, given the structure of the regulator to be tuned, it is quite simple to figure out what the phase of the process frequency response could be in the vicinity of the cutoff frequency. A complete discussion could not fit here, suffice to say that in practical every case that is tractable by a PI(D) autotuner, the "good" cutoff frequencies are those for which the process frequency response phase is in the range $(-180°, 90°)$. Therefore, some frequency domain information in that band, that is easily found e.g. by relay feedback tests, is a good starting point here.

In the following, it will be shown that such information is very well suited to derive methods that comprehend both the model identification and the subsequent tuning, and (which is a strength of the proposal) can seamlessly be based on existing model-based tuning rules (although new ones could be devised within the same framework).

III. Model-based tuning with contextual model parametrisation

This section presents the general concept of the "model-based tuning with contextual model parametrisation" methodology, termed in the following the "contextual method" for brevity. That method is based on the idea that it makes no sense to separate the parameterisation of the model from the subsequent tuning of the regulator, treating the overall model-based tuning process as the solution of two *cascaded* algebraic systems. The core of the methodology, simple but extremely efficient, is to treat the tuning rules and the model parameterisation rules *jointly*, and therefore to solve a *single* algebraic system.

The methodology, at the generality level of this treatise, needs the following "ingredients":

1) a regulator structure
$$R(s, \theta_R) \qquad \theta_R \in \Re^{n_R} \tag{1}$$
where θ_R is the vector of regulator parameters;
2) a process model structure
$$M(s, \theta_M) \qquad \theta_M \in \Re^{n_M} \tag{2}$$
where θ_M is the vector of the model parameters;
3) a tuning rule capable of determining the parameters of (1) based on those of (2) and on n_D design variables forming a vector $\theta_D \in \Re^{n_D}$, i.e., n_T equations that can be expressed in the form
$$g_T(\theta_M, \theta_R, \theta_D) = 0 \tag{3}$$
or more frequently
$$\theta_R = f_T(\theta_M, \theta_D) \tag{4}$$
since in general explicit tuning rules are sought and devised, and in that case (the only one considered in the examples reported later on) clearly $n_T = n_R$;
4) n_P points of the process Nyquist curve $P(j\omega_i)$, $i = 1 \ldots n_P$, found e.g. with relay experiment(s), although the method used to determine those point is completely irrelevant for the purpose of the presented research.

Hence, the problem of parametrising M and tuning R so far has $n_R + n_M + n_D$ variables (the regulator parameters, the model parameters, and the design variables of the tuning rule) and n_T equations (those substantiating the tuning rule itself, recall that for explicit tuning rules $n_T = n_R$).

Contrary to the classical model-based tuning procedure, where first M is found and then R is tuned after deciding somehow θ_D, the idea here is to put everything together in a single system of equations. To this end, first write that "the model is exact at the known points of the process frequency response", i.e.,
$$M(j\omega_i, \theta_M) = P(j\omega_i) \qquad \forall i = 1 \ldots n_P, \tag{5}$$
which provides $2n_P$ more real equations to the problem.

Then, take the chosen model-based tuning rule, employ it to express the nominal cutoff frequency ω_{cn} of the control system containing the tuned regulator and the tuning model - which is possible by definition with any such rule, although details are omitted here for brevity - and write one more equation saying that ω_{cn} equals one of the frequencies ω_i of the known points of the process frequency response.

At this point, the overall problem has $n_T + 2n_P + 1$ real equations. To make it solvable, it is therefore enough to add n_f real equations, where n_f is such that
$$n_R + n_M + n_D = n_T + 2n_P + 1 + n_f. \tag{6}$$

Two are the key points of this reasoning. First, the model is found *together* with the regulator tuning, and is by construction "exact at the cutoff frequency". Second, and more important here, the added n_f equations can fix design variables, regulator parameters, model parameters or any combination thereof, or even simply impose some relationship between those quantities (for example, thinking of a real PID, one could fix the high-frequency control sensitivity by constraining the quantity $K(1+N)$ to a desired value).

In synthesis, under the sole constraint that the obtained system of equations is mathematically tractable (and a few examples follow to show that this happens in many interesting cases) there is here no distinction, no hierarchy among the various sets of variables stemming from their role in the problem: everything in the process of parametrising the model and tuning the regulator here is treated jointly, everything is in one word "contextual".

IV. Some applications of the contextual method

The contextual method can be applied in principle to any model-based tuning rule. To witness that, this section presents a few examples of such applications, referring to well known model-based PI tuning rules.

A. IMC contextual method fixing λ

The first example deals with the IMC-PI rules [16], [8], that refer to the model structure

$$M(s) = \frac{\mu}{1+sT} e^{-sL} \tag{7}$$

and compute the PI parameters as

$$T_i = T, \qquad K = \frac{T}{\mu(L+\lambda)} \tag{8}$$

where *traditionally* λ is interpreted as the desired closed-loop time constant.

With the contextual method, one can start from a single point

$$P(\overline{\omega}) = A_P e^{j\varphi_P} \tag{9}$$

of the process Nyquist curve, require that the model frequency response contain that point, i.e.,

$$M(\overline{\omega}) = P(\overline{\omega}) \tag{10}$$

and impose that the nominal cutoff frequency

$$\omega_{cn} = \frac{1}{L+\lambda}. \tag{11}$$

equal $\overline{\omega}$. This yields the system

$$\begin{cases} T_i &= T \\ K &= \frac{T}{\mu(L+\lambda)} \\ A_P &= \frac{\mu}{\sqrt{1+(\bar{\omega}T)^2}} \\ \varphi_P &= -\arctan(\bar{\omega}T) - \bar{\omega}L \\ \overline{\omega} &= \frac{1}{L+\lambda} \end{cases} \tag{12}$$

having 5 equations and 6 unknowns. If λ is fixed, one obtains

$$\begin{cases} L &= \frac{1}{\bar{\omega}} - \lambda \\ T &= -\frac{1}{\bar{\omega}} tan(\bar{\omega}L + \varphi_P) \\ \mu &= A_P\sqrt{1+(\bar{\omega}T)^2} \\ T_i &= T \\ K &= \frac{T}{\mu(L+\lambda)} \end{cases} \tag{13}$$

which, by setting

$$\theta'_D = [\lambda], \quad \theta'_M = [\mu\, T\, L], \quad \theta'_R = [K\, T_i] \tag{14}$$

corresponds to fulfilling the balance (6) with $n_P = 1$, $n_T = n_R = 2$, $n_M = 3$, $n_D = 1$, and $n_f = 1$.

B. IMC contextual method fixing K

One could reason in the same way as section IV-A, but solve (12) fixing the PI high-frequency magnitude K, thus the high-frequency control sensitivity, instead of λ. The equations/variables balance (6) clearly still holds, and one obtains

$$\begin{cases} T &= \frac{KA_P}{\bar{\omega}\sqrt{1+(KA_P)^2}}; \\ T_i &= T; \\ \mu &= \frac{T\bar{\omega}}{K}; \\ L &= -\frac{1}{\bar{\omega}}(arctan(\bar{\omega}T) + \varphi_P); \\ \lambda &= \frac{T}{K\mu} - L. \end{cases} \tag{15}$$

V. Symmetric Optimum method fixing τ

The Symmetric Optimum (SO) tuning rule [17], [18] for a PI takes as process model

$$M(s) = \frac{\mu}{1+sT} e^{-sL} \prod_{h=1}^{n} (1+sT_k) \tag{16}$$

i.e., a FOPDT augmented with some additional poles accounting for unmodelled dynamics. Defining

$$T_{um} = L + \tau \tag{17}$$

where

$$\tau := \sum_{h=1}^{n} (1+sT_h) \tag{18}$$

as a quantification of those unmodelled dynamics (τ taking traditionally the meaning of "amount of unmodelled *rational* dynamics" under the hypothesis of a "realistic" delay estimate), the tuning rule is

$$T_i = 4T_{um}, \qquad K = \frac{T}{2\mu T_{um}}. \tag{19}$$

and (18) has to be counted as a tuning relationship, hence in this particular case $n_T = 3$.

Starting again from a point of the process Nyquist curve, see (9), recalling that the nominal cutoff frequency is here

$$\omega_c = \frac{1}{2T_{um}} \tag{20}$$

and reasoning in the same way as with the IMC-PI rule yields the system

$$\begin{cases} T_i &= 4T_{um} \\ K &= \frac{T}{2\mu T_{um}} \\ A_P &= \frac{\mu}{\sqrt{1+(\bar{\omega}T)^2}\sqrt{1+(\bar{\omega}T_{um})^2}} \\ \varphi_P &= -\arctan(\bar{\omega}T) - \arctan(\bar{\omega}T_{um}) - \bar{\omega}L \\ \bar{\omega} &= \frac{1}{2T_{um}} \\ T_{um} &= L + \tau \end{cases} \tag{21}$$

with 6 equations and 7 unknowns. Now, if τ is fixed,

$$\begin{cases} T_{um} &= \frac{1}{2\bar{\omega}} \\ L &= T_{um} - \tau \\ T &= -\frac{1}{\bar{\omega}} tan(arctan(\frac{1}{2}) + \bar{\omega}L + \varphi_P) \\ \mu &= \frac{\sqrt{5}}{2} A_P \sqrt{1+(\bar{\omega}T)^2} \\ K &= \frac{\bar{\omega}T}{\mu} \\ T_i &= \frac{2}{\bar{\omega}} \end{cases} \tag{22}$$

is obtained, which by setting

$$\theta'_D = [\tau], \quad \theta'_M = [\mu\, T\, L\, T_{um}], \quad \theta'_R = [K\, T_i] \tag{23}$$

corresponds to fulfilling (6) with $n_P = 1$, $n_T = 3$, $n_R = 2$, $n_M = 4$, $n_D = 1$, and $n_f = 1$.

VI. Symmetric Optimum method fixing K

System (21) can also be solved after fixing the high-frequency PI gain K, obtaining

$$\begin{cases} T_{um} &= \frac{1}{2\bar{\omega}} \\ T_i &= \frac{2}{\bar{\omega}} \\ \mu &= A_P \frac{\frac{\sqrt{5}}{2}}{\sqrt{1-A_P\frac{5}{4}K^2}} \\ T &= K\frac{\mu}{\bar{\omega}} \\ L &= -\frac{1}{\bar{\omega}}(arctan(\bar{\omega}T)+arctan(\bar{\omega}T_{um})+\varphi_P) \\ \tau &= L-T_{um}. \end{cases} \tag{24}$$

VII. CDS contextual method fixing λ

The Chen-Seborg tuning method [19], often called the "CDS" method in the literature, can synthesise a PI with a FOPDT model by the tuning rule

$$T_i = \frac{T^2+TL-(\lambda-T)^2}{T+L}, \quad K = \frac{T^2+TL-(\lambda-T)^2}{(\lambda+L)^2} \tag{25}$$

where λ is interpreted in an analogous way to the IMC-PI. Omitting lengthy computations, the nominal cutoff frequency is here

$$\omega_c = \frac{T+L}{(\lambda+L)^2} \tag{26}$$

and the usual reasoning based on one point of the process Nyquist curve leads to the system

$$\begin{cases} T_i &= \frac{T^2+TL-(\lambda-T)^2}{T+L} \\ K &= \frac{1}{\mu}\frac{T^2+TL-(\lambda-T)^2}{(\lambda+L)^2} \\ A_P &= \frac{\mu}{\sqrt{1+(\bar{\omega}T)^2}} \\ \varphi_P &= -\arctan(\bar{\omega}T)-\bar{\omega}L \\ \bar{\omega} &= \frac{T+L}{(\lambda+L)^2} \end{cases} \tag{27}$$

with 5 equations and 6 unknowns. Fixing parameter λ gives

$$\begin{cases} L &= -\frac{1}{\bar{\omega}}(arctan(\bar{\omega}^2(\lambda+L)^2-\bar{\omega}L)+\varphi_P) \\ T &= \bar{\omega}(L+\lambda)^2-L \\ \mu &= A_P\sqrt{1+(\bar{\omega}T)^2} \\ T_i &= \frac{T^2+TL-(\lambda-T)^2}{T+L} \\ K &= \frac{1}{\mu}\frac{T^2+TL-(\lambda-T)^2}{(\lambda+L)^2} \end{cases} \tag{28}$$

that with

$$\theta_D' = [\lambda], \quad \theta_M' = [\mu\, T\, L], \quad \theta_R' = [K\, T_i] \tag{29}$$

means fulfilling (6) with $n_P = 1$, $n_T = n_R = 2$, $n_M = 3$, $n_D = 1$, and $n_f = 1$.

A. Some conclusions on the contextual method application

Many other examples could be reported, but doing so would add little information to this manuscript. The few ones of the previous sections should be enough to prove that the contextual method can be used with virtually any tuning rule available in the literature, which is definitely a strength point.

VIII. A comparative example

This section analyses the contextual method through a comparison between the results that can be obtained with it and with classical approaches to the tuning model identification.

To set up the comparison, first we chose two classical identification method, namely the method of areas and the tangent method. Then, we chose two model based PI tuning methods, the IMC and the CDS. Those methods were chosen because they are different enough, in that they are based on different tuning approaches, but have also relevant similarities, because they use the same design variable (λ) and they require similar specifications on the system, As such, the resulting comparison is meaningful, and can be interpreted in a sensible way.

The comparison is made on a batch of four processes, namely

$$P_1(s) = \frac{1}{(1+s)(1+5s)}, \quad P_2(s) = \frac{(1-0.3s)}{(1+s)^3},$$
$$P_3(s) = \frac{1}{(1+2\cdot\frac{0.6}{1}s+s^2)}, \quad P_4(s) = \frac{1+s}{(1+2s)(1+0.2s)^2}. \tag{30}$$

For a better comparison, when not using the contextual method, λ was invariantly selected as $0.5/\omega_{90}$, ω_{90} being the frequency at which the frequency response of each process has phase $-90°$. The results are reported in figures 1 through 4. Each figure is composed of six plots, that report the response of the controlled variable to a unit load disturbance step as forecast with the FOPDT model (in blue) and as obtained with the real process (in red). Going from left to right and from top to bottom, the six plots refer to the results of

- the IMC-PI tuning rules with the FOPDT model found with the method of areas,
- the IMC-PI tuning rules with the FOPDT model found with the tangent method,
- the contextual tuning method based on the IMC-PI tuning rules,
- the CDS PI tuning rules with the FOPDT model found with the method of areas,
- the CDS PI tuning rules with the FOPDT model found with the tangent method,
- the contextual tuning method based on the CDS PI tuning rules.

With the simple, overdamped process $P1$, the method of areas coupled to the IMC-PI rule gives quite good results, but the settling time is larger than with the contextual method based on the same rule. As for the tangent method, with both the IMC-PI and the CDS rules, the results are definitely worse. Notice that the transients forecast with the model are reasonably correct with the method of areas and the contextual method, and definitely erratic with the tangent method.

The situation is very similar with the non-minimum phase process P_2, while things change a bit for the loosely damped process P_3, where it is the method of areas that produces the

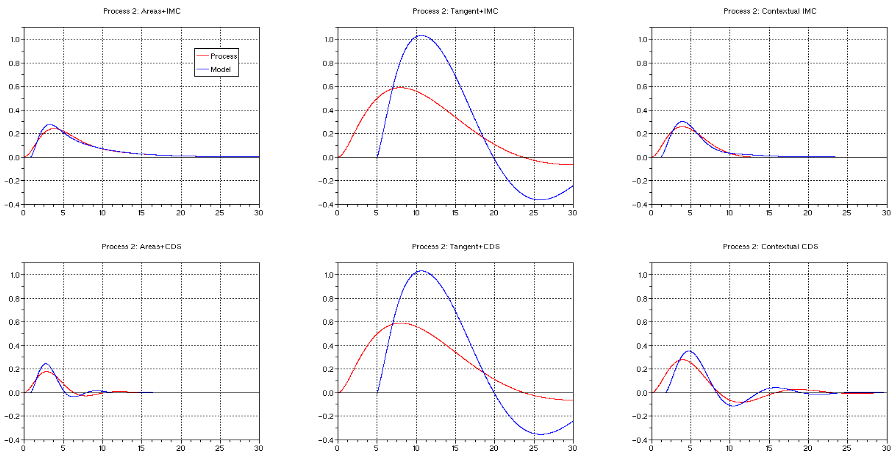

Fig. 1. Comparison results with process P_1.

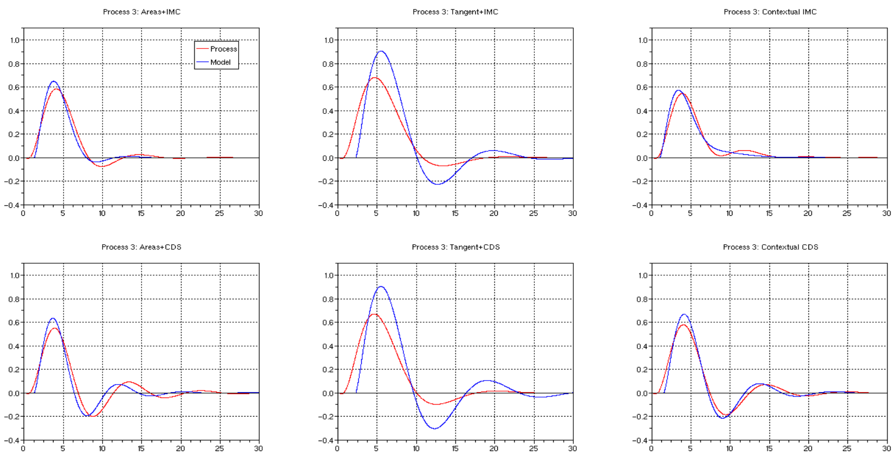

Fig. 2. Comparison results with process P_2.

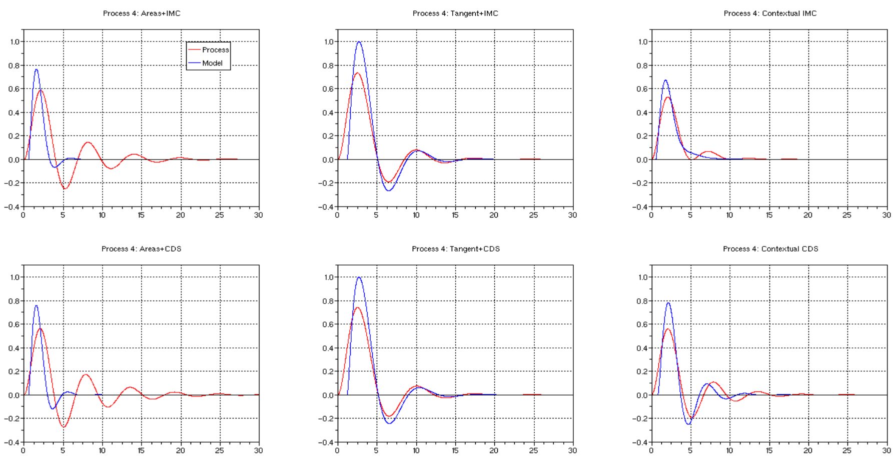

Fig. 3. Comparison results with process P_3.

worst results, and also fails at providing a model capable of forecasting the closed-loop transients correctly.

So far, it was just shown that the contextual method performs reasonably when applied to well-established tuning rules, and that having a model that is by construction precise near the cutoff frequency invariantly improves the capability of that model to forecast the closed-loop transients.

Process P_4, that does not exhibit a particularly complex

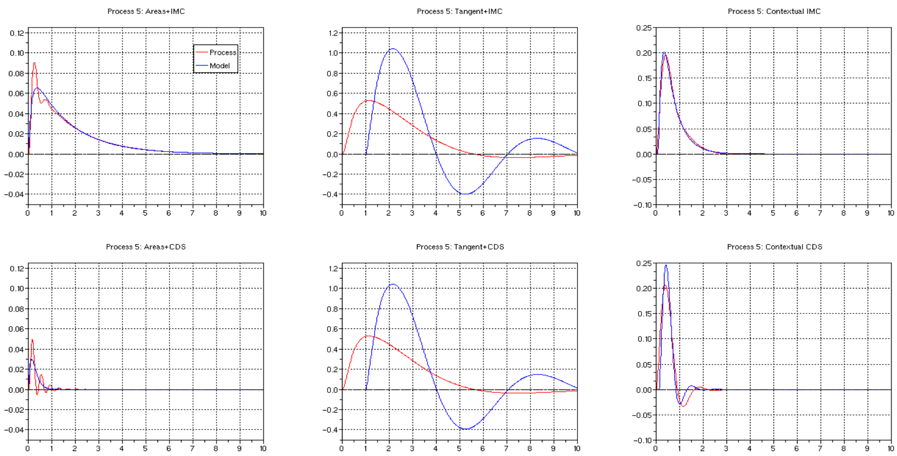

Fig. 4. Comparison results with process P_4.

dynamics but has a zero and a pole at near frequencies (a characteristic frequently encountered in process control, especially with thermal systems) best highlights here the advantages of the contextual method. As can be seen, the results obtained with the method of areas and the two tuning rules are completely different, and the model forecasts are quite poor, especially with the CDS rule. In addition, the actual degree of stability is not stunning, and the control sensitivity is quite large, particularly with the CDS rule (we omit figures on that for brevity, but the fact is evident based on the depicted transients). Conversely, the tangent method with both rules provides results that are acceptable and similar to one another, but the model forecasts are even worse than with the method of areas. In this case, the contextual method not only allows both rules to provide better results, with a good compromise among response speed, stability and control effort, but also yields models that are capable of forecasting those results reliably.

IX. Conclusions and future work

A methodology was presented to complement any existing model-based PI(D) tuning rule with a "contextual" model identification, based on frequency domain data.

The overall result are complete tuning procedures, that provide good control results, and also process models capable of forecasting those results reliably. And above all, such procedures do not separate model identification from regulator tuning, which is a very significant advantage.

The focus was restricted here to the PI case, but the idea is general. It can be extended to the PID, and also to other regulator structures. It can be used with existing rules, allowing by the way to establish meaningful relationships and comparisons among them, and also to create new ones. Research is underway on all of those subjects, and the results will be presented in future works.

References

[1] A. Leva, "Model-based tuning: the very basics and some useful techniques," *Journal-A*, vol. 42, no. 3, pp. 14–22, 2001.

[2] A. Datta, M. Ho, and S. Bhattacharyya, *Structure and synthesis of PID controllers*. Springer-Verlag, 2000.

[3] C. Brosilow and B. Joseph, *Techniques of model based control*. Indianapolis, IN: Prentice Hall PTR, 2002.

[4] K. Åström, T., and Hägglund, *Advanced PID Control*. Research Triangle Park, NY: The Instrumentation, Systems, and Automation Society, 2005.

[5] A. O'Dwyer, *Handbook of PI and PID controller tuning rules*. Singapore: World Scientific Publishing, 2003.

[6] A. Isaksson and S. Graebe, "Analytical PID parameter expression for higher order systems," *Automatica*, vol. 35, pp. 1121–1130, 1999.

[7] S. Skogestad, "Simple analytic rules for model reduction and PID controller tuning," *Journal of Process Control*, vol. 13, pp. 291–309, 2003.

[8] A. Leva and A. Colombo, "On the IMC-based synthesis of the feedback block of ISA-PID regulators," *Transactions of the Institute of Measurement and Control*, vol. 26, no. 5, pp. 417–440, 2004.

[9] A. Leva and F. Schiavo, "On the role of the process model in model-based autotuning," in *Proc. 16th IFAC World Congress*, Praha, Czech Republic, 2005.

[10] S. Bittanti, M. Campi, and S. Garatti, "New results on the asymptotic theory of system identification for the assessment of the quality of estimated models," in *Proc. 41st CDC*, Las Vegas, NV, 2002, pp. 1814–1819.

[11] M. Gevers, "Identification for control: from the early achievements to the revival of experiment design," *European Journal of Control*, vol. 11, no. 4–5, pp. 335–352, 2005.

[12] H. Hjalmarsson, "From experiment design to closed-loop control," *Automatica*, vol. 43, pp. 393–438, 2005.

[13] A. Leva, "Simple model-based PID autotuners with rapid relay identification," in *Proc. 16th IFAC World Congress*, Praha, Czech Republic, 2005.

[14] A. Leva, L. Bascetta, and F. Schiavo, "Model-based Proportional-Integral/Proportional-Integral-Derivative (PI/PID) autotuning with fast relay identification," *Industrial & Engineering Chemistry Research*, vol. 45, no. 7, pp. 4052–4062, 2006.

[15] A. Leva and L. Piroddi, "On the parametrisation of simple process models for the autotuning of industrial regulators," in *Proc. 1ACC 2007*, New York, NY, 2007.

[16] R. Braatz, "Internal model control," in *The control handbook*, S. Levine, Ed. Boca Raton, FL: CRC Press, 1996, pp. 215–224.

[17] C. Kessler, "Das symmetrische optimum, teil I," *Regelungstechnik*, vol. 6, pp. 432–436, 1958.

[18] ——, "Das symmetrische optimum, teil II," *Regelungstechnik*, vol. 6, pp. 395–400, 1958.

[19] D. Chen and D. Seborg, "PI/PID controller design based on direct synthesis and disturbance rejection," *Ind. Eng. Chem. Process Des. Dev.*, vol. 41, pp. 4807–4822, 2002.