

PI/PID autotuning with contextual model parametrisation

Alberto Leva,*

Sara Negro, Alessandro Vittorio Papadopoulos¹

*Dipartimento di Elettronica e Informazione, Politecnico di Milano
Via Ponzio 34/5, 20133 Milano, Italy*

Abstract

In model-based PI/PID tuning regulators, the same set of I/O data and the same tuning rule can produce very different results, depending just on the procedure used to parametrise the process model. The problem is seldom addressed, but extremely relevant for the acceptability of model-based autotuners in the applications. This manuscript proposes a methodology to treat the model parametrisation and regulator tuning phases *jointly*, so as to circumvent said problem with affordable process stimulation and computational effort. The methodology can be generalised to different regulator structures, and even employed to devise new tuning rules.

Key words: Autotuning; model-based tuning; PID control.

*Corresponding author: voice (39) 02 2399 3410, fax (39) 02 2399 3412.

Email addresses: leva@elet.polimi.it (Alberto Leva.)

¹Graduate students at the Dipartimento di Elettronica e Informazione.

1. Introduction and motivation

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, which is *subsequently* employed - together with convenient specifications - to compute the regulator parameters with some “tuning rules”. A review of the huge MBAT literature is impossible to give here, the interested reader can refer e.g. to [10, 17, 7, 4], and for a broad *panorama* to the excellent survey [30].

In general, MBAT methods are based on very simple models, of structure decided *a priori* based essentially on that of the regulator, despite the process dynamics encountered may be complex [1], and require a “clever” order reduction [34]. Two are the main reasons for the fact above. First, models need identifying on-line, based on data produced by *stimuli* that must obey to potentially severe process upset constraints and therefore typically lack excitation, which hampers e.g. model order selection. Second, to achieve tuning procedures suitable for an industrial implementation, *explicit* tuning rules are desirable [13, 31], which call for a simple model unconditionally.

Such a *scenario* motivates the widespread use in MBAT of *ad hoc* parametrisation methods such as that of areas, of moments, of the tangent, and so forth. As a consequence of the heuristics unavoidably introduced, the results obtained with MBAT depend on the used parametrisation method *significantly*. The problem is seldom addressed in the literature, but hampers a wide acceptance of MBAT in the application domain. In fact, if the same set of data, the same model structure and the same tuning rules produce different tuning results depending just on which method is selected to find the model parameters, it is not surprising that the industrial community’s confidence in MBAT at large is adversely affected, as easily

observed in technology reviews such as [28].

Indeed, in MBAT there is hardly any point in discussing tuning rules without taking into account the effects of the model parametrisation method [23, 27]. This is however complex: in the light of the above remarks, on one hand well established results of the identification theory often prove inadequate to assess the quality of a model *for the purpose of MBAT*. On the other hand, process *stimuli* limitations leave limited or no room for experiment design, so that MBAT can take little (if any) profit of the neat products of the “identification for control” research [11, 12]. In fact, virtually the totality of MBAT methods could be termed “serial”, in that they comprehend two distinct phases: *first* the model is determined, *then* the regulator is tuned. In general little attention is paid to model assessment or validation, and even if said issue is somehow addressed, this is obviously done without using any information on the (subsequent) tuning.

The main contribution of this work, that is part of a long-term research initiated by [19] and continued by [26], is the definition of a methodology to complement any existing MBAT rule with a contextual procedure to parametrise the required model, whence the adjective “contextual” applied to the proposed method as a counterpart to the “serial” character of the others noted above. In the authors’ opinion, this is a step forward for at least two reasons. First, joining the model parametrisation and the tuning phases circumvents the mentioned problem, in that for any process and tuning rules, the proposed procedure invariantly yields results comparable, if not superior, to those produced by the parametrisation method that with the same process and rules performs best among those classically used. Second, the obtained model is by construction precise near the cutoff frequency, hence providing sensible and reliable forecasts of the closed-loop transients, which is *not*

guaranteed with typical heuristic parametrisations. It is also worth noticing that in this research only *stimuli* that can be applied to real-world processes are considered.

The focus is here restricted to the PI/PID regulator, and to the use of frequency-domain data—namely, measured points of the process Nyquist curve. However, the idea as expressed in section 2 and used in section 3 is far more general.

2. Model-based tuning with contextual model parametrisation

This section presents the general idea of the “model-based tuning with contextual model parametrisation” methodology, termed in the following the “contextual method” for brevity. The adjective “contextual” indicates precisely that there is no distinction here between the model parametrisation and the regulator tuning phase, the most distinctive feature of the method being exactly that the two corresponding systems of equations, irrespectively of the particular tuning rule used, are treated and solved as a *single* system.

2.1. Foreword

Two are the key points of this research: (a) join the parametrisation and tuning phases, and (b) in doing so use only process *stimuli* suitable for real-world industrial implementations. A very important problem is then which type of process information to employ within the set that point (b) permits, and before entering the proposed method’s explanation, some words on the matter are useful.

Seeking information in the time domain (e.g., the observed delay, time constant and gain of a step response, very frequent in the applications) leads to catch essentially the low-frequency process dynamics, and taking that dynamics *a priori* as the “control-relevant” one may be *very* misleading [27]. Note also that most

MBAT rule operate, more or less implicitly, by cancellation, whence for example (unduly) general statements such as “MBAT rules give good tracking but poor load disturbance rejection”. One must bear in mind that the *model*, not the “process’ poles are cancelled. If the model is a low-frequency approximation of the process, cancelling its poles will most likely lead to excessively low-frequency regulator zeroes, and *that* is the cause for the mentioned problem.

On the other hand, given the regulator structure, it is simple to figure out what the phase of the process frequency response could be near the cutoff frequency. Suffice here to say that in virtually every case tractable by a PI/PID, the “good” cutoff frequencies are those for which said phase is in the range $(-180^\circ, -90^\circ)$. Therefore, some frequency domain information in that band, that is easily found e.g. by relay feedback tests, is a good answer to the problem of choosing which type of process data to use (at least for the purpose of this study); note, incidentally, that the idea is consistent with recent research trends [16]. Such information is therefore used in this work, although in principle different answers could be found for the same question, abstracting the following proposal from its *declination* into the use of relay tests.

2.2. The method

At the generality level of this treatise, and based on the above considerations, the contextual method refers to

1. a regulator structure, expressed in the Laplace transform domain as

$$R(s, \theta_R) \quad \theta_R \in \mathfrak{R}^{n_R} \quad (1)$$

where θ_R is a vector of parameters;

2. a process model (structure)

$$M(s, \theta_M) \quad \theta_M \in \mathfrak{R}^{n_M} \quad (2)$$

where θ_M is a vector of 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 $\theta_D \in \mathfrak{R}^{n_D}$, i.e., n_T equations in the form $g_T(\theta_M, \theta_R, \theta_D) = 0$, or $\theta_R = f_T(\theta_M, \theta_D)$ in *explicit* tuning rules, where clearly $n_T = n_R$;
4. n_P points of the process Nyquist curve $P(j\omega_i)$, $i = 1 \dots n_P$, found e.g. with relay experiment(s), although the method used to determine those points is irrelevant for this research.

Hence, the problem of parametrising M and tuning R with *explicit* rules (the only case treated here for brevity) so far has $n_R + n_M + n_D$ variables.

Contrary to the classical MBAT procedure, the corresponding $n_R + n_M + n_D$ equations are here tested all together. To this end, first write that “the model is exact at the known points of the process frequency response”, i.e.,

$$M(j\omega_i, \theta_M) = P(j\omega_i) \quad \forall i = 1 \dots n_P, \quad (3)$$

which provides $2n_P$ real equations. Then, use the chosen tuning rules to express the nominal cutoff frequency ω_{cn} of the control system - which is possible by construction with any such rule, details are omitted for brevity - and write one more equation saying that ω_{cn} equals one of the frequencies ω_i of the known frequency response points.

At this point, the overall problem has $n_T + 2n_P + 1$ real equations. It is therefore enough to add n_f real equations, where

$$n_R + n_M + n_D = n_T + 2n_P + 1 + n_f. \quad (4)$$

Two are the key points of this reasoning. First, not only is the model found *together* with the regulator tuning, but it is by construction “exact at the cutoff frequency”. Second, and more important, the added n_f equations can prescribe the values of design variables, regulator parameters, model parameters or any combination thereof, or even simply impose some relationship between those quantities. In synthesis, then, under the sole constraint that the obtained system of equations be 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 based on their role in the overall tuning problem: everything in the process of parametrising the model and tuning the regulator here is treated jointly, everything is in one word “contextual”.

3. Some applications of the contextual method

This section presents a few applications of the contextual method to well known model-based PI/PID tuning rules, to show that the method can be used with *any* such rule.

3.1. IMC PI with prescribed λ

The first example deals with the IMC-PI rules [6, 23], that refer to the model structure

$$M(s) = \frac{\mu}{1 + sT} e^{-sL} \quad (5)$$

and compute the parameters of the PI

$$R_{PI}(s) = K \left(1 + \frac{1}{sT_i} \right) \quad (6)$$

as

$$T_i = T, \quad K = \frac{T}{\mu(L + \lambda)} \quad (7)$$

where *traditionally* λ is interpreted as the desired closed-loop time constant, although it has significant relationships with the achieved robustness [22]. With the contextual method, one can start from a single point

$$P(\bar{\omega}) = A_P e^{j\varphi_P} \quad (8)$$

of the process Nyquist curve, require that the model frequency response contain that point, i.e., $M(\bar{\omega}) = P(\bar{\omega})$, and impose that the nominal cutoff frequency $\omega_{cn} = 1/(L + \lambda)$ equal $\bar{\omega}$. This yields the system

$$\begin{cases} T_i = T \\ K = \frac{T}{\mu(L+\lambda)} \\ A_P = \frac{\mu}{\sqrt{1+(\bar{\omega}T)^2}} \\ \varphi_P = -\arctan(\bar{\omega}T) - \bar{\omega}L \\ \bar{\omega} = \frac{1}{L+\lambda} \end{cases} \quad (9)$$

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

$$\begin{cases} L = \frac{1}{\bar{\omega}} - \lambda \\ T = -\frac{1}{\bar{\omega}} \tan(\bar{\omega}L + \varphi_P) \\ \mu = A_P \sqrt{1+(\bar{\omega}T)^2} \\ T_i = T \\ K = \frac{T}{\mu(L+\lambda)} \end{cases} \quad (10)$$

which, by setting

$$\theta'_D = [\lambda], \quad \theta'_M = [\mu T L], \quad \theta'_R = [K T_i] \quad (11)$$

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

3.2. IMC PI with prescribed high-frequency control sensitivity

One could reason in the same way as section 3.1, but solve (9) fixing K , thus the high-frequency control sensitivity, instead of λ . The equations/variables balance (4) clearly still holds, and one obtains

$$\left\{ \begin{array}{l} T = \frac{KA_P}{\bar{\omega}\sqrt{1+(KA_P)^2}}; \\ T_i = T; \\ \mu = \frac{T\bar{\omega}}{K}; \\ L = -\frac{1}{\bar{\omega}}(\arctan(\bar{\omega}T) + \varphi_P); \\ \lambda = \frac{T}{K\mu} - L. \end{array} \right. \quad (12)$$

3.3. Symmetric Optimum PI with prescribed τ

The Symmetric Optimum (SO) tuning rule [14, 15] for a PI takes as process model

$$M(s) = \frac{\mu}{1+sT} e^{-sL} \left(\prod_{h=1}^n (1+sT_h) \right)^{-1} \quad (13)$$

Defining $T_{um} = L + \tau$ where

$$\tau := \sum_{h=1}^n (1+sT_h) \quad (14)$$

takes traditionally the meaning of “amount of unmodelled (rational) dynamics”, the tuning rules are

$$T_i = 4T_{um}, \quad K = \frac{T}{2\mu T_{um}}. \quad (15)$$

and (14) has to be counted as a tuning relationship, hence in this case $n_T = 3$. Starting again from a point of the process Nyquist curve, see (8), recalling that the nominal cutoff frequency is here $\omega_{cn} = 1/2T_{um}$, and reasoning in the same way as

with the IMC-PI rule, one obtains the system

$$\left\{ \begin{array}{l} T_i = 4T_{um} \\ K = \frac{T}{2\mu T_{um}} \\ A_P = \frac{\mu}{\sqrt{1+(\bar{\omega}T)^2}\sqrt{1+(\bar{\omega}T_{um})^2}} \\ \varphi_P = -\arctan(\bar{\omega}T) - \arctan(\bar{\omega}T_{um}) - \bar{\omega}L \\ \bar{\omega} = \frac{1}{2T_{um}} \\ T_{um} = L + \tau \end{array} \right. \quad (16)$$

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

$$\left\{ \begin{array}{l} T_{um} = \frac{1}{2\bar{\omega}} \\ L = T_{um} - \tau \\ T = -\frac{1}{\bar{\omega}}\tan(\arctan(\frac{1}{2}) + \bar{\omega}L + \varphi_P) \\ \mu = \frac{\sqrt{5}}{2}A_P\sqrt{1+(\bar{\omega}T)^2} \\ K = \frac{\bar{\omega}T}{\mu} \\ T_i = \frac{2}{\bar{\omega}} \end{array} \right. \quad (17)$$

which by setting

$$\theta'_D = [\tau], \quad \theta'_M = [\mu T L T_{um}], \quad \theta'_R = [K T_i] \quad (18)$$

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

3.4. Symmetric Optimum PI with prescribed high-frequency control sensitivity

Also (16) can be solved after fixing K , which yields

$$\left\{ \begin{array}{l} T_{um} = \frac{1}{2\bar{\omega}} \\ T_i = \frac{2}{\bar{\omega}} \\ \mu = A_P \frac{\frac{\sqrt{5}}{2}}{\sqrt{1 - A_P \frac{5}{4} K^2}} \\ T = K \frac{\mu}{\bar{\omega}} \\ L = -\frac{1}{\bar{\omega}} (\arctan(\bar{\omega}T) + \arctan(\bar{\omega}T_{um}) + \varphi_P) \\ \tau = L - T_{um}. \end{array} \right. \quad (19)$$

3.5. CDS PI with prescribed λ

The ‘‘CDS’’ tuning method [9] can synthesise a PI with a model in the form (5) by the rules

$$T_i = \frac{T^2 + TL - (\lambda - T)^2}{T + L}, \quad K = \frac{1}{\mu} \frac{T^2 + TL - (\lambda - T)^2}{(\lambda + L)^2} \quad (20)$$

where λ is interpreted like in IMC-PI. The nominal cutoff frequency is here $\omega_{cn} = (T + L)(\lambda + L)^2$, and the usual reasoning based on one frequency response point yields

$$\left\{ \begin{array}{l} T_i = \frac{T^2 + TL - (\lambda - T)^2}{T + L} \\ K = \frac{1}{\mu} \frac{T^2 + TL - (\lambda - T)^2}{(\lambda + L)^2} \\ A_P = \frac{\mu}{\sqrt{1 + (\bar{\omega}T)^2}} \\ \varphi_P = -\arctan(\bar{\omega}T) - \bar{\omega}L \\ \bar{\omega} = \frac{T + L}{(\lambda + L)^2} \end{array} \right. \quad (21)$$

with 5 equations and 6 unknowns. Fixing λ gives

$$\begin{cases} L = -\frac{1}{\bar{\omega}}(\arctan(\bar{\omega}^2(\lambda + L)^2 - \bar{\omega}L) + \varphi_P) \\ T = \bar{\omega}(L + \lambda)^2 - L \\ \mu = A_P\sqrt{1 + (\bar{\omega}T)^2} \\ T_i = \frac{T^2 + TL - (\lambda - T)^2}{T + L} \\ K = \frac{1}{\mu} \frac{T^2 + TL - (\lambda - T)^2}{(\lambda + L)^2} \end{cases} \quad (22)$$

that with

$$\theta'_D = [\lambda], \quad \theta'_M = [\mu \ T \ L], \quad \theta'_R = [K \ T_i] \quad (23)$$

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

3.6. IMC PID with prescribed λ

The last two applications refer to the PID structure. Referring again to (5), the IMC-PID rules [23] compute the parameters of the real (ISA) PID

$$R_{PID}(s) = K \left(1 + \frac{1}{sT_i} + \frac{sT_d}{1 + sT_d/N} \right) \quad (24)$$

with

$$T_i = T + \frac{L^2}{2(L + \lambda)}, \quad K = \frac{T_i}{\mu(L + \lambda)}, \quad N = \frac{T(L + \lambda)}{\lambda T_i} - 1, \quad T_d = \frac{\lambda LN}{2(L + \lambda)} \quad (25)$$

where λ has the same meaning as in section 3.1. Using the contextual method with one frequency response point $P(\bar{\omega}) = A_P e^{j\varphi_P}$, requiring (5) to be exact at $\bar{\omega}$,

and that $\omega_{cn} = 1/(L + \lambda)$ equal $\bar{\omega}$, one obtains the system

$$\left\{ \begin{array}{l} T_i = T + \frac{L^2}{2(L+\lambda)} \\ K = \frac{T_i}{\mu(L+\lambda)} \\ N = \frac{T(L+\lambda)}{\lambda T_i} - 1 \\ T_d = \frac{\lambda LN}{2(L+\lambda)} \\ A_P = \frac{\mu}{\sqrt{1+(\bar{\omega}T)^2}} \\ \varphi_P = -\arctan(\bar{\omega}T) - \bar{\omega}L \\ \bar{\omega} = \frac{1}{L+\lambda} \end{array} \right. \quad (26)$$

having 7 equations in 6 unknowns. Here fixing λ leads to

$$\left\{ \begin{array}{l} L = -\lambda + \frac{1}{\bar{\omega}} \\ T = -\frac{1}{\bar{\omega}} \tan(\bar{\omega}L + \varphi_P) \\ \mu = A_P \sqrt{1+(\bar{\omega}T)^2} \\ T_i = T + \frac{L^2}{2(L+\lambda)} \\ K = \frac{T_i}{\mu(L+\lambda)} \\ N = \frac{T(L+\lambda)}{\lambda T_i} - 1 \\ T_d = \frac{\lambda LN}{2(L+\lambda)} \end{array} \right. \quad (27)$$

which, by setting

$$\theta'_D = [\lambda], \quad \theta'_M = [\mu T L], \quad \theta'_R = [K T_i N T_d] \quad (28)$$

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

3.7. CDS PID with prescribed λ and high-frequency control sensitivity

Also the CDS method [9] has a PID version, namely the rule

$$\begin{aligned}
 K &= \frac{-2\lambda^3 - 3L\lambda^2 + \left(2LT + \frac{L^2}{2}\right) \left(3\lambda + \frac{L}{2}\right)}{\mu \left(2\lambda^3 + 3L\lambda^2 + \frac{L^2(3\lambda + \frac{L}{2})}{2}\right)} \\
 T_i &= \frac{-2\lambda^3 - 3L\lambda^2 + \left(2LT + \frac{L^2}{2}\right) \left(3\lambda + \frac{L}{2}\right)}{L(2T + L)} \\
 T_d &= \frac{-2(T + L)\lambda^3 + 3LT\lambda^2 + \frac{L^2T(3\lambda + \frac{L}{2})}{2}}{-2\lambda^3 - 3L\lambda^2 + \left(2LT + \frac{L^2}{2}\right) \left(3\lambda + \frac{L}{2}\right)}
 \end{aligned} \tag{29}$$

that, notice, refers (contrary to the IMC-PID one) to an *ideal* PID. If however one relates the parameters of the ideal PID (marked with a tilde) to those of the real one (24), the relationship

$$\begin{cases}
 T_i + \frac{T_d}{N} &= \tilde{T}_i \\
 T_i T_d \left(1 + \frac{1}{N}\right) &= \tilde{T}_i \tilde{T}_d \\
 \frac{K}{T_i} &= \frac{\tilde{K}}{\tilde{T}_i} \\
 K(1 + N) &= R_\infty
 \end{cases} \tag{30}$$

is obtained, where R_∞ is the high-frequency control sensitivity magnitude with the real PID. The nominal cutoff frequency is here

$$\omega_{cn} = \frac{8LT + 4L^2}{8\lambda^3 + 12L\lambda^2 + 6L^2\lambda + L^3} \tag{31}$$

and based again on one frequency response point, one obtains to a system with 10 equations and 12 unknowns (omitted for brevity), that solved after fixing both R_∞

and λ yields

$$\left\{ \begin{array}{l} L = -\frac{1}{\bar{\omega}}(\text{atan}(\bar{\omega}^2(\lambda + L)^2 - \bar{\omega}L) + \varphi_P) \\ T = \frac{\bar{\omega}(8\lambda^3 + 12L\lambda^2 + 6L^2\lambda + L^3)}{8L} - \frac{L}{2} \\ \mu = A_P \sqrt{\bar{\omega}^2 T^2 + 1} \\ \tilde{K} = \frac{-2\lambda^3 - 3L\lambda^2 + (2LT + \frac{L^2}{2})(3\lambda + \frac{L}{2})}{\mu \left(2\lambda^3 + 3L\lambda^2 + \frac{L^2(3\lambda + \frac{L}{2})}{2} \right)} \\ \tilde{T}_i = \frac{-2\lambda^3 - 3L\lambda^2 + (2LT + \frac{L^2}{2})(3\lambda + \frac{L}{2})}{L(2T + L)} \\ \tilde{T}_d = \frac{-2(T + L)\lambda^3 + 3LT\lambda^2 + \frac{L^2 T(3\lambda + \frac{L}{2})}{2}}{-2\lambda^3 - 3L\lambda^2 + (2LT + \frac{L^2}{2})(3\lambda + \frac{L}{2})} \\ K = \frac{\tilde{K}(R_\infty \tilde{T}_i - \tilde{K} \tilde{T}_d)}{R_\infty \tilde{T}_i} \\ N = \frac{\tilde{K}^2 \tilde{T}_d - \tilde{K} R_\infty \tilde{T}_i + R_\infty^2 \tilde{T}_i}{\tilde{K}(R_\infty \tilde{T}_i - \tilde{K} \tilde{T}_d)} \\ T_d = \frac{\tilde{T}_d(\tilde{K}^2 \tilde{T}_d - \tilde{K} R_\infty \tilde{T}_i + R_\infty^2 \tilde{T}_i)}{R_\infty(R_\infty \tilde{T}_i - \tilde{K} \tilde{T}_d)} \\ T_i = \frac{R_\infty \tilde{T}_i - \tilde{K} \tilde{T}_d}{R_\infty} \end{array} \right. \quad (32)$$

and with

$$\theta'_D = [\lambda \ R_\infty], \quad \theta'_M = [\mu \ T \ L], \quad \theta'_R = [\tilde{K} \ \tilde{T}_i \ \tilde{T}_d \ K \ T_i \ T_d \ N] \quad (33)$$

means fulfilling (4) with $n_P = 1$, $n_T = n_R = 7$, $n_M = 3$, $n_D = 2$, and $n_f = 2$.

4. Implementation considerations

With respect to some autotuners in the literature, the proposed procedures seem to disregard facts like the model order suggested by the used data, and the type of control problem (e.g., set point tracking versus disturbance rejection) to be addressed—both surely of relevance for the applicability of the method(s). According to the mainstream autotuning literature, however (see e.g. [30]), structured information on the model that best represents the observed I/O data is useful,

but basically just to possibly tailor the rule to be used, see e.g. the discussion in [25]. Considering also the well known difficulties in designing reliable tuning rules for high-order models [13], *in the specific domain of PID autotuning* the interest is better focused on “how a *simple* model can fit the data” than on “what is the best order to use”. Indeed, in the autotuning literature the model structure is hardly ever dictated by anything else than the *regulator* structure [5]. In some sense, the quest for a tuning made with the *minimum* possible information has a very successful representative in the numerous techniques based on relay feedback [33]. The presented research tries to break the idea of relay-based tuning in its “identification” (of one or more points) and “tuning” phases, to observe that the latter involves no ambiguity and no information loss, and thus to use the former as a possible solution to the model parametrisation problem as declined in autotuning.

As for the type of control problem, one may object that the idea of “finding a model intrinsically exact at the cutoff frequency” constrains the joint parametrisation/tuning procedure so that no degrees of freedom may be left to manage e.g. a tradeoff among load disturbance recovery speed, high-frequency control sensitivity, and stability degree. In fact, however, what is proposed here is to employ parametrisation methods that privilege a narrow frequency band (the relay case being the extremum where a single frequency is considered) and to use model based tuning rules setting a frequency in that band as the required (nominal) cut-off. The above possible concern is therefore answered by driving the parametrisation phase so as to centre the model on a given band, and possibly exploiting the two-degree-of-freedom structures of industrial regulators.

In synthesis, then, the proposed (contextual) procedure can be summarised as

follows.

1. Choose *a priori* a regulator structure, represented by vector $\theta_R \in \mathfrak{R}^{n_R}$.
2. Choose *a priori* a model structure, represented by vector $\theta_M \in \mathfrak{R}^{n_M}$.
3. Define the required process information (here one or more points of the process Nyquist curve).
4. Find a set of tuning rules for the selected regulator structure with respect to the selected model structure.
5. Condition the model so that it catches the process information exactly (here it must be exact in the used points of the Nyquist curve).
6. Write a system of equation imposing the facts above, plus the required specification(s) that must hold exactly if the previous point holds (here the cutoff frequency is made equal to one of the frequencies of the points).
7. Solve (symbolically or numerically) the obtained system, finding in one step the regulator and the tuning model.

It is worth noticing that one could apply the proposed idea to a wide class of methods, but not all the possible groups of methods allow for a meaningful comparison of the achieved results. For example, methods that one wants to compare should share at least the model structure and possibly the set of design variables, or at least allow for a re-interpretation of said variables with a common meaning, so that it be possible to “ask all the methods to do the same thing”. Needless to say, comparable methods have also to share the regulator structure. This is the case for the IMC and the CDS, while other rules such as for example the AMIGO [5] refer to an ideal PID, and others like the Ziegler-Nichols ones are conceived devoid of design variables.

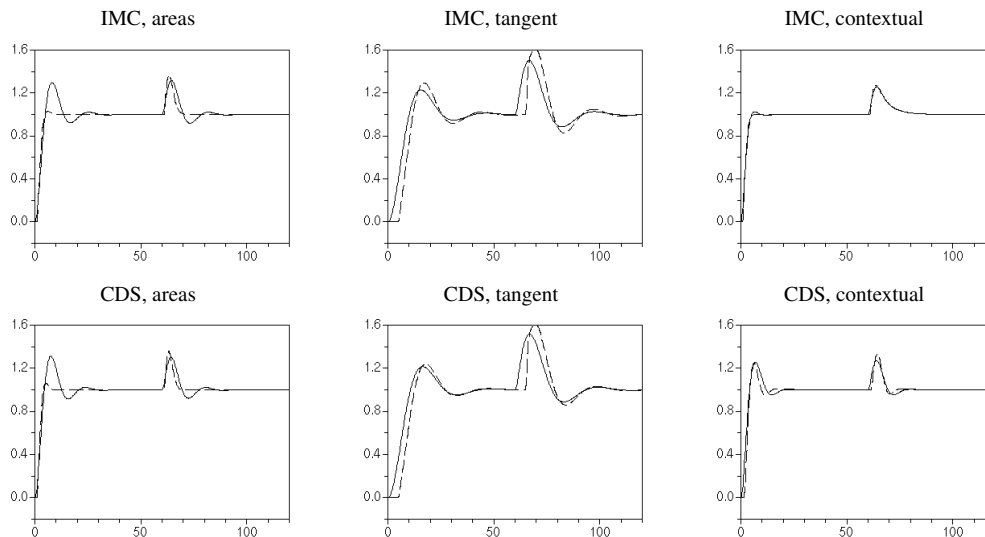


Figure 1: Test with process P_1 - set point and load disturbance step responses of the controlled variable with the process (solid) and as forecast with the model (dashed).

5. Three comparative examples

5.1. Example 1 (PI)

This example analyses the contextual method for PI rules through a comparison between the results that can be obtained with it and with classical approaches. The comparison uses two identification methods, namely the method of areas and the tangent method, and two model based PI tuning methods, the IMC and the CDS. Those methods were chosen because they are based on different tuning approaches, but use the same design variable (λ) and require similar specifications. As such, the resulting comparison is meaningful, and can be interpreted

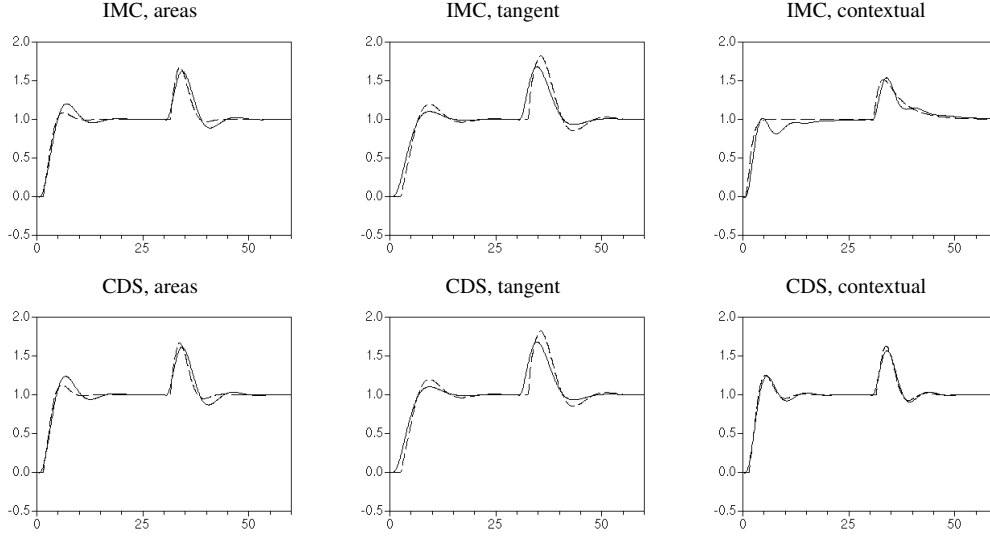


Figure 2: Test with process P_2 - set point and load disturbance step responses of the controlled variable with the process (solid) and as forecast with the model (dashed).

in a sensible way. The comparison is made on a batch of four processes, namely

$$\begin{aligned}
 P_1(s) &= \frac{1}{(1+s)(1+5s)}, & P_2(s) &= \frac{(1-0.3s)}{(1+s)^3}, \\
 P_3(s) &= \frac{1}{(1+1.2s+s^2)}, & P_4(s) &= \frac{1+s}{(1+2s)(1+0.2s)^2}.
 \end{aligned} \tag{34}$$

designed along a reasoning similar to that of [3]. 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 set point and load disturbance step as obtained with the process, and as forecast with the model. Going from left to right and from top to bottom, the six plots refer to the results of (a) the IMC-PI tuning rules with the model found with the method of areas,

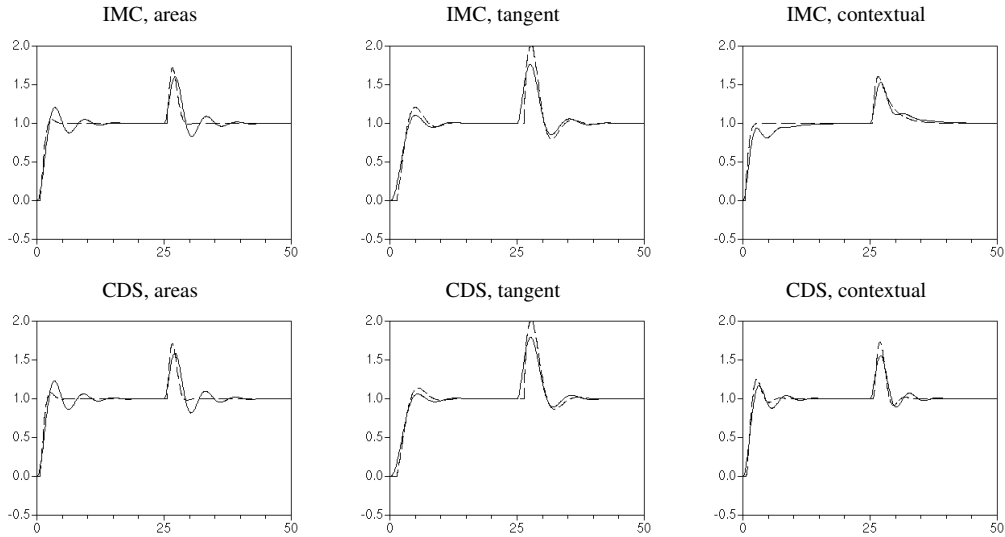


Figure 3: Test with process P_3 - set point and load disturbance step responses of the controlled variable with the process (solid) and as forecast with the model (dashed).

(b) the IMC-PI tuning rules with the model found with the tangent method, (c) the contextual tuning method based on the IMC-PI tuning rules, (d) the CDS PI tuning rules with the model found with the method of areas, (e) the CDS PI tuning rules with the model found with the tangent method, and (f) the contextual tuning method based on the CDS PI tuning rules. The reported comparisons are based on intuitive considerations: of course one may want to introduce some objective indicators for that purpose, for example based on the considerations of [24], but doing so would have been lengthy here, and possibly distract the reader from the major remarks that follow.

With the simple, overdamped process P_1 , 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

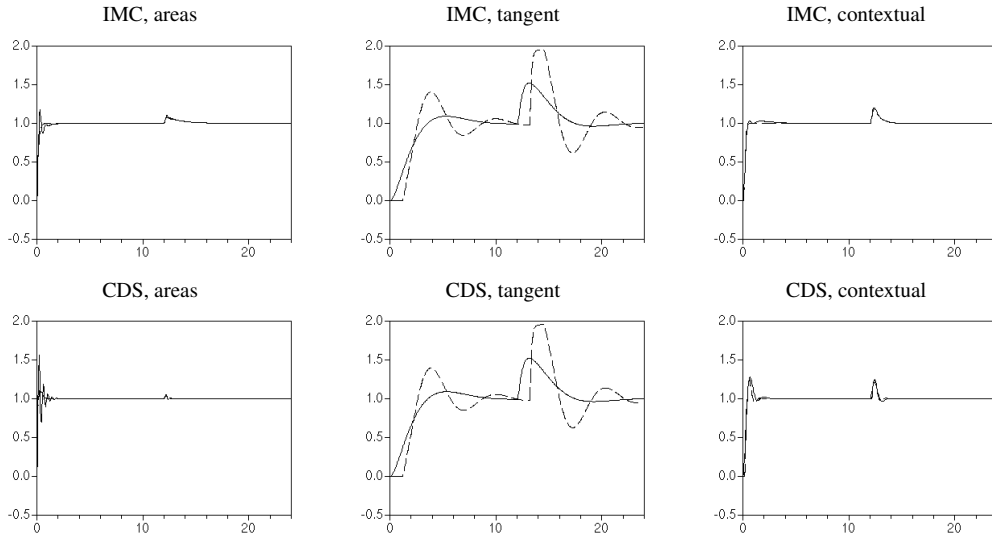


Figure 4: Test with process P_4 - set point and load disturbance step responses of the controlled variable with the process (solid) and as forecast with the model (dashed).

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

Process P_4 further highlights here the advantages of the contextual method. 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 rules. 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

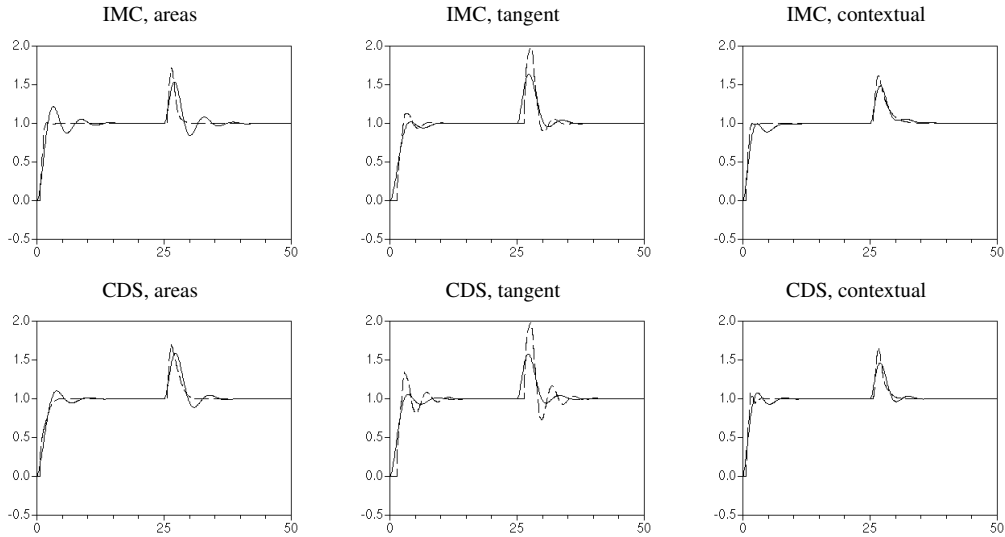


Figure 5: Test with process P_3 - set point and load disturbance step responses of the controlled variable with the process (solid) and as forecast with the model (dashed).

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. For example the method of areas yields better disturbance rejection but far more nervous a set point response, thus a control action. Also, the contextual method yields models that are capable of forecasting those results reliably.

5.2. Example 2 (PID)

This example is analogous to that of section 5.1, but refers to the two PID rules the contextual method was applied to. Again, the areas and tangent methods are used. The comparison is made on the process $P_3(s)$, the same process used previously. When not using the contextual method, λ is selected as $0.6/\omega_{90}$, ω_{90} being the frequency at which the frequency response of each process has phase

		ISE M		ISE P		ISD M/P	
		SP	LD	SP	LD	SP	LD
P_1	A	1.993	0.382	2.399	1.429	0.393	0.091
PI	T	7.940	2.673	4.619	3.278	0.998	0.898
	C	1.698	0.365	1.529	0.941	0.016	0.002
P_2	A	2.445	1.265	2.568	1.955	0.067	0.083
PI	T	4.028	2.595	3.057	2.489	0.142	0.319
	C	1.477	1.146	2.055	1.538	0.206	0.062
P_3	A	1.056	0.673	1.175	0.945	0.091	0.139
PI	T	2.180	2.072	1.694	1.529	0.069	0.244
	C	0.757	0.805	1.002	0.815	0.144	0.040
P_4	A	0.066	0.006	0.087	0.047	0.018	$2 \cdot 10^{-4}$
PI	T	1.978	1.731	1.130	0.839	0.410	0.976
	C	0.167	0.019	0.159	0.089	0.003	10^{-4}
P_3	A	0.867	0.519	1.001	0.768	0.154	0.174
PID	T	1.713	1.466	1.217	1.039	0.113	0.286
	C	0.821	0.568	0.817	0.637	0.072	0.031

Table 1: Integral indices relative to the simulation tests with P_1 – P_4 and the IMC rules.

-90° . With the CDS method, that refers to an ideal PID, parameter N in the used ISA form was set to 10 (a typical value in practice). The results are reported in figure 5. Apparently, the method of areas coupled to the IMC-PID rules fails at providing a model capable of forecasting the closed-loop transients correctly. Also the method of areas with the CDS rules produces inferior results with respect to the contextual method based on the same rules. The tangent method, with both the IMC-PI and the CDS rules, does not provide good results, both in terms of stability degree and forecast transients. On the other hand, with both rules the contextual method shows comparable, uniform results.

To end this set of examples, table 1 reports some integral indices (explained in the legend) relative to the simulation tests, so as to quantify both the tuning quality and the fidelity of the model forecasts. The first column indicates the used

process and regulator type, the second column indicates the areas (**A**), tangent (**T**) or contextual (**C**) method. The labels on the first two rows are interpreted as follows: **ISE M, SP** is the Integral of the Squared Error (set point minus controlled variable) for a unit step set point variation computed with the model, **ISE P, SP** is the same as ISE M, SP but computed with the process, **ISE M, LD** and **ISE P, LD** are the same as ISE M, SP and ISE P, SP but for a unit step load disturbance, **ISD M/P, SP** is the Integral of the Squared Difference between the controlled variable as predicted with the model and as obtained with the process for a unit step set point variation, **ISD M/P, LD** is the same as ISD M/P, SP but for a unit step load disturbance.

In any vertical group of three indices, the bold one is the best (smallest) value. It can be seen that the contextual method is superior for both the SP and LD cases except for P_4 , where however the obtained results (figure 4) are comparable to those of the method of areas, indicating that integral indices are significant, but cannot replace completely the inspection of transients. Note anyway that with P_4 the contextual method provides the best model forecasts. Finally, the only cases in which the method of areas provides better forecasts (P_2 and P_3 with a PI) are relative to set point tracking, where the use of a two-degree-of-freedom (2-dof) structure, that is not within the scope of this work although some words on it are spent later on, may have an influence.

In synthesis, once again, the proposed method gives comparable or better results, with particular reference to the quality of the closed loop transients' forecasts.

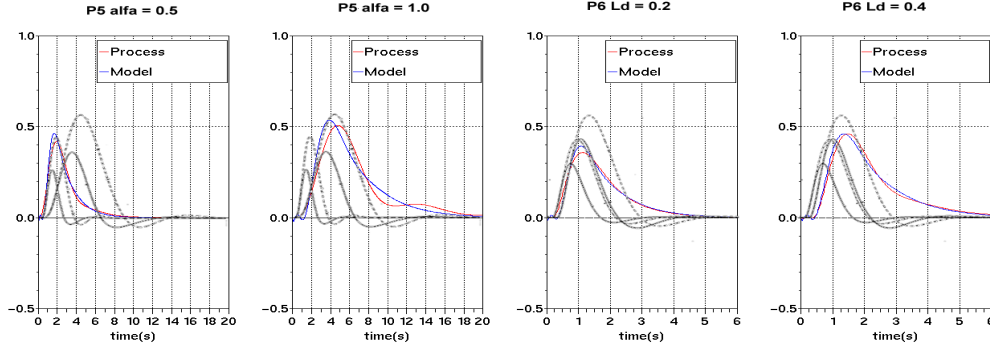


Figure 6: Comparison with the procedure of [16], processes P_5 and P_6 with various α and L_d . Black lines are the results of the quoted reference.

5.3. Example 3 (comparison with frequency domain tuning)

This section compares the proposed PID procedure with that reported in [16], that employs selected frequency response points, and can be taken as a good, recent representative of tuning methods grounded in the frequency domain (based e.g. on relay experiments). The processes considered in the following are

$$P_5(s) = \frac{1}{(1+s)(1+\alpha s)(1+\alpha^2 s)(1+\alpha^3 s)}, \quad \alpha = 0.5, 1.0 \quad (35)$$

and

$$P_6(s) = \frac{e^{-sL_d}}{(1+s)(0.2s)}, \quad L_d = 0.2, 0.4 \quad (36)$$

that correspond, respectively, to (12) and (13) in section 6.1 of [16], where the interested reader can find full detail. Figure 6 shows the controlled variable's load disturbance step responses obtained in the four cases with the contextual method, and the IMC rules. The results from [16] are reported in the figure for comparison. Incidentally, in this case the method of areas performs quite well, while the tangent method is generally inferior.

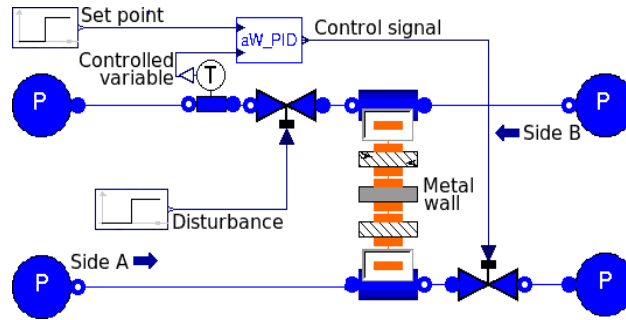


Figure 7: Modelica diagram of the model used in the simulation example.

6. A more complex example

As a final test, we consider quite realistic a situation, namely an accurate, non-linear model of a counterflow heat exchanger, including mass, energy and momentum equations, friction and heat exchange correlations, and detailed water/steam properties calculation based on the IF97 standard tables [32]. The model (having about 1600 equations) was implemented in the Modelica language [29, 2], by using component models from the ThermoPower library [8], and its scheme is shown in figure 7 (the full Modelica code can be obtained by contacting the corresponding author). An accurate simulator was preferred to a physical test since comparisons among different rules call for highly repeatable conditions, which are difficult to guarantee with any experimental setup.

The objective is to control the output temperature of the (heated) side B by acting on the valve of the (heating) side A, the side B valve introducing a disturbance. A step test, followed by the application of the areas and the tangent method, and a relay test providing the -90° point for the contextual method, yield the regulators of table 2 (only the IMC-PID rule was used for brevity, structuring the comparison as in section 5.2). Figure 8 compares the so obtained PIDs, the

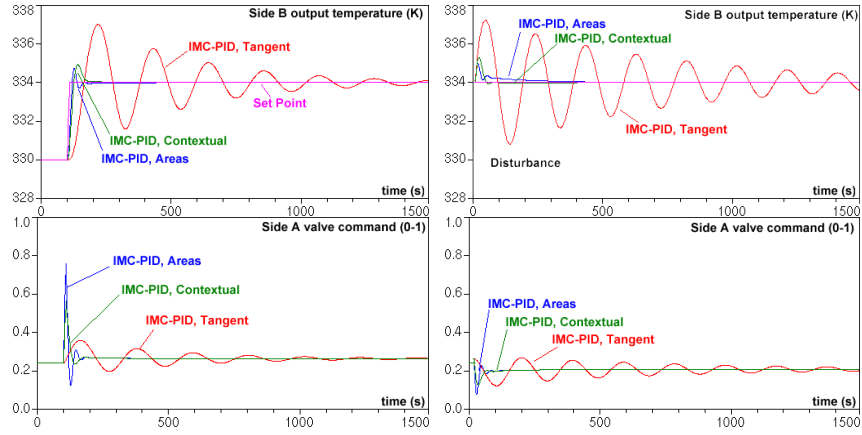


Figure 8: Simulation results - responses of the controlled variable (top) and the control signal (bottom) to a side B output temperature set point quick ramp (4°K in 200s, left) and to a side B valve step from 0.8 to 0.3 (right).

latter concentrating on the response to the disturbance introduced by an abrupt, step-like closing of the side B valve.

IMC-PID	K	T_i	T_d	N
Method of Areas	0.18	235	0.05	0.11
Tangent method	0.0008	1.05	0.026	0.06
Contextual tuning	0.081	36.4	1.09	0.61

Table 2: Regulator parameters in the simulation example.

The basically delay-free aspect of the step response makes the use of the tangent method quite critical, while the method of areas causes the IMC-PID to set T_i to the (excessively large) estimated time constant, see table 2. The method of areas and the contextual one give comparable results for set point tracking. The larger overshoot of the contextual method is an indicator of a stronger feedback in the control band, and can anyway be eliminated by a suitably tuned 2-dof struc-

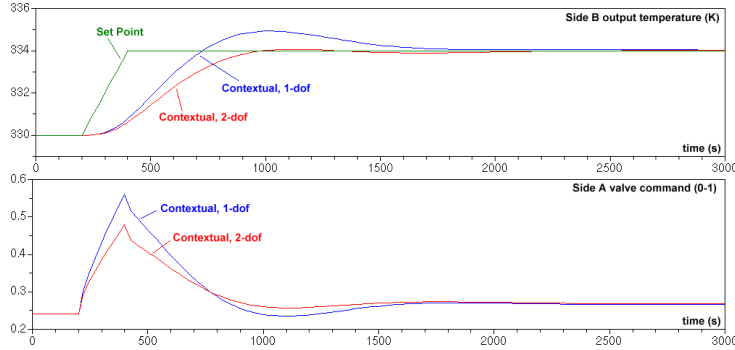


Figure 9: Simulation results - comparison of the results obtained with a 1-dof and a 2-dof PID.

ture. For example, one may use the PID procedure of [20]. Said procedure tunes the feedforward path (i.e., the set point weights b and c in the proportional and derivative action) of the 2-dof ISA PID expressed as

$$U(s) = K \left(bW(s) - Y(s) + \frac{1}{sT_i}(W(s) - Y(s)) + \frac{sT_d}{1 + sT_d/N}(cW(s) - Y(s)) \right) \quad (37)$$

where $W(s)$, $Y(s)$ and $U(s)$ are, respectively, the Laplace transforms of the set point, the controlled variable, and the control signal. The used method accepts as input the other PID parameters (i.e., those of the 1-dof version) and some responses of the 1-dof control loop, termed the "base functions", that can be obtained easily if a reliable process model is available—see [20] for details impossible to give here. The quoted method is therefore very suited for use in conjunction with the contextual one presented here. If the tracking specification (the dominant time constant of the set point response) is set to 2, the quoted method produces $b = 0.68$ and $c = 0.32$. The obtained set point responses are presented in figure 9, compared with the 1-dof ones shown previously.

As for disturbance rejection, the lower high-frequency control sensitivity magnitude (0.13 versus 0.2) of the contextual method results in a slightly larger initial

deviation of the controlled variable from the set point, but the recovery time is definitely improved.

7. Conclusions and future work

A methodology was presented to complement any existing model-based PI/PID 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 the system behaviour. 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/PID case, but the idea is general, and can be extended to other regulator structures, either similar to the PID but conceived for particular tuning goals, e.g. [18], or even totally different, provided a model-based procedure is available for them, and the systems of equations playing the same role of those seen here turn out to be tractable. The method can be used with all existing rules, allowing by the way to establish meaningful relationships and comparisons among them, and also to create new ones, for example by further exploiting frequency domain data to estimate model error overbounds and therefore enhance robustness [21]. Research is underway on all of those subjects, and the results will be presented in future works.

References

- [1] J. Alvarez-Ramirez and R. Femat. Robust PI stabilization of a class of chemical reactors. *Systems & Control Letters*, 38:219–225, 1999.
- [2] Modelica Association. Modelica home page.

- [3] K.J. Åström and T. Hägglund. Benchmark systems for PID control. In *IFAC Workshop on Digital Control – Past, present, and future of PID Control*, Terrassa, Spain, 2000.
- [4] K.J. Åström, T., and Hägglund. *Advanced PID Control*. The Instrumentation, Systems, and Automation Society, Research Triangle Park, NY, 2005.
- [5] K.J. Åström, T., and Hägglund. *Advanced PID Control*. The Instrumentation, Systems, and Automation Society, Research Triangle Park, NY, 2006.
- [6] R.D. Braatz. Internal model control. In S. Levine, editor, *The control handbook*, pages 215–224. CRC Press, Boca Raton, FL, 1996.
- [7] C. Brosilow and B. Joseph. *Techniques of model based control*. Prentice Hall PTR, Indianapolis, IN, 2002.
- [8] F. Casella and A. Leva. Modelling of thermo-hydraulic power generation processes using Modelica. *Mathematical and Computer Modelling of Dynamical Systems*, 12(1):19–33, 2006.
- [9] D. Chen and D.E. Seborg. PI/PID controller design based on direct synthesis and disturbance rejection. *Ind. Eng. Chem. Process Des. Dev.*, 41:4807–4822, 2002.
- [10] A. Datta, M. Ho, and S.P. Bhattacharyya. *Structure and synthesis of PID controllers*. Springer-Verlag, 2000.
- [11] M. Gevers. Identification for control: from the early achievements to the revival of experiment design. *European Journal of Control*, 11(4–5):335–352, 2005.

- [12] H. Hjalmarsson. From experiment design to closed-loop control. *Automatica*, 43:393–438, 2005.
- [13] A.J. Isaksson and S.F. Graebe. Analytical PID parameter expression for higher order systems. *Automatica*, 35:1121–1130, 1999.
- [14] C. Kessler. Das symmetrische optimum, teil I. *Regelungstechnik*, 6:432–436, 1958.
- [15] C. Kessler. Das symmetrische optimum, teil II. *Regelungstechnik*, 6:395–400, 1958.
- [16] B. Kristiansson and B. Lennartson. Evaluation and simple tuning of PID controllers with high-frequency robustness. *Journal of Process Control*, 16(2):91–102, 2006.
- [17] A. Leva. Model-based tuning: the very basics and some useful techniques. *Journal-A*, 42(3):14–22, 2001.
- [18] A. Leva. Autotuning process controller with improved load disturbance rejection. *Journal of Process Control*, 15(2):223–234, 2005.
- [19] 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*, 45(7):4052–4062, 2006.
- [20] A. Leva and A.M. Colombo. Method for optimising set-point weights in ISA-PID autotuners. *IEE Proceedings - Control Theory and Applications*, 146(2):137–146, 1999.

- [21] A. Leva and A.M. Colombo. Estimating model mismatch overbounds for the robust autotuning of industrial regulators. *Automatica*, 36(12):1855–1861, 2000.
- [22] A. Leva and A.M. Colombo. Implementation of a robust PID autotuner in a control design environment. *Transactions of the Institute of Measurement and Control*, 23(1):1–20, 2001.
- [23] A. Leva and A.M. Colombo. On the IMC-based synthesis of the feedback block of ISA-PID regulators. *Transactions of the Institute of Measurement and Control*, 26(5):417–440, 2004.
- [24] A. Leva and F. Donida. Quality indices for the autotuning of industrial regulators. *IET Control Theory & Applications*, 3(21):170–180, 2009.
- [25] A. Leva and L. Piroddi. Model classification applied to industrial PID autotuning. *Industrial & Engineering Chemistry Research*, 43(21):6787–6796, 2004.
- [26] A. Leva and L. Piroddi. On the parametrisation of simple process models for the autotuning of industrial regulators. In *Proc. IACC 2007*, New York, NY, 2007.
- [27] 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.
- [28] Y. Li, K.H. Ang, and C.Y. Chong. Patents, software, and hardware for PID control—an overview and analysis of the current art. *IEEE Control Systems Magazine*, pages 42–54, february 2006.

- [29] S.E. Mattsson, H. Elmqvist, and M. Otter. Physical system modeling with Modelica. *Control Engineering Practice*, 6(4):501–510, 1998.
- [30] A. O’Dwyer. *Handbook of PI and PID controller tuning rules*. World Scientific Publishing, Singapore, 2003.
- [31] S. Skogestad. Simple analytic rules for model reduction and PID controller tuning. *Journal of Process Control*, 13:291–309, 2003.
- [32] W. Wagner, J. R. Cooper, A. Dittmann, J. Kijima, H.-J. Kretzschmar, A. Kruse, R. Mareš, K. Oguchi, H. Sato, I. Stöcker, O. Šifner, Y. Takaishi, I. Tanishita, J. Trübenbach, and Th. Willkommen. The iapws industrial formulation 1997 for the thermodynamic properties of water and steam. *Journal of Engineering for Gas Turbines and Power*, 122(1):150–184, 2000.
- [33] C.C. Yu. *Autotuning of PID Controllers: Relay Feedback Approach*. Springer Verlag, London, 1999.
- [34] K. Zhou, C. D’Souza, and J.R. Cloutier. Structurally balanced controller order reduction with guaranteed closed loop performance. *Systems and Control Letters*, 24:235–242, 1995.

A. Numerical data

	μ	L	T	K	T_i
Area + IMC-PI	0.600012	1.1398487	1.723803	1.1531317	1.6754212
Tangent + IMC-PI	0.600012	5.16	0.84	0.2149972	0.84
Contextual IMC-PI	0.9643342	0.9010546	5.3279863	2.4526998	5.3279863
Area + CDS-PI	0.600012	1.1398487	1.723803	1.2882101	1.6754212
Tangent + CDS-PI	0.600012	5.16	0.84	0.1878188	0.7963807
Contextual CDS-PI	0.5507794	1.690019	2.4168735	1.7252652	2.1405463

Table 3: Comparison with process P_1 .

	μ	L	T	K	T_i
Area + IMC-PI	0.9055629	1.5182625	1.389306	0.5697829	1.389306
Tangent + IMC-PI	0.9055629	2.62	1.18	0.3434226	1.18
Contextual IMC-PI	1.8345443	0.7828835	4.629234	1.2892697	4.629234
Area + CDS-PI	0.9055629	1.5182625	1.389306	0.6082359	1.3734107
Tangent + CDS-PI	0.9055629	2.62	1.18	0.3439337	1.1799915
Contextual IMC-PI	1.0477997	1.4683771	2.0999065	0.9068913	1.8598189

Table 4: Comparison with process P_2 .

	μ	L	T	K	T_i
Area + IMC-PI	1.0299432	0.6331983	0.6943163	0.5457826	0.6943163
Tangent + IMC-PI	1.0299432	1.42	0.34	0.1632646	0.34
Contextual IMC-PI	2.1468593	0.40131	2.3729687	1.1017128	2.3729687
Area + CDS-PI	1.0299432	0.6331983	0.6943163	0.5811621	0.6878917
Tangent + CDS-PI	1.0299432	1.42	0.34	0.1258144	0.3010081
Contextual CDS-PI	1.2261783	0.7526975	1.0764226	0.774961	0.9533525

Table 5: Comparison with process P_3 .

	μ	L	T	K	T_i
Area + IMC-PI	0.9639615	0	1.3225964	10.342921	1.3225964
Tangent + IMC-PI	0.9639615	1.24	0.12	0.0906901	0.12
Contextual IMC-PI	0.7188783	0.0884368	0.522932	3.2901567	0.522932
Area + CDS-PI	0.9639615	0	1.3225964	19.648455	0.2520053
Tangent + CDS-PI	0.9639615	1.24	0.12	0.0897658	0.1198822
Contextual CDS-PI	0.4105873	0.1658722	0.2372117	2.3143447	0.2100907

Table 6: Comparison with process P_4 .

	μ	L	T	K	T_i	T_d	N
Area+IMC	1	0.6455078	0.5544922	0.5783709	0.7215020	0.0923003	0.5926405
Tangent+IMC	1	1.42	0.34	0.4147569	0.8386239	0.0764764	0.3618024
Contextual-IMC	2.1468593	0.4013100	2.3729687	1.1389766	2.4532307	0.0736972	0.6121386
Area+CDS	1	0.6455078	0.5544922	0.5243958	0.7322603	0.2972848	10
Tangent+CDS	1	1.42	0.34	0.6611291	1.0013258	0.1220271	10
Contextual-CDS	1.1030785	0.7526975	0.8635723	1.0653097	1.1789688	0.2432696	10

Table 7: Comparison with process P_3 controlled by a PID.

	μ	L	T	K	T_i	T_d	N
$\alpha = 0.5$	1.1164495	0.628367	1.5308025	1.4704614	1.7193126	0.1540622	1.2258932
$\alpha = 1.0$	1.2895661	1.4471483	3.5254846	1.2730607	3.9596291	0.3548098	1.2258932

Table 8: Simulations referring to process P_5 .

	μ	L	T	K	T_i	T_d	N
$L_d = 0.2$	0.9454375	0.4024373	0.9894916	1.7364405	1.1011328	0.098669	1.2258932
$L_d = 0.4$	1.1088398	0.5044637	1.2289541	1.4805528	1.3802932	0.1236837	1.2258932

Table 9: Simulations referring to process P_6 .