Proceeding of the 2004 American Control Conference
Boston, Massachusetts June 30 - July 2, 2004

Autotuning process controller with enhanced load disturbance rejection

Alberto Leva
Dipartimento di Elettronica e Informazione
Politecnico di Milano
Via Ponzio, 34/5 – 20133 Milano, Italy
leva@elet.polimi.it

Abstract – **This paper presents an autotuning process controller aimed at providing efficient rejection of load disturbances, in a class of situations that are typical in process cotnrol, and not easy to treat with most standard autotuning controllers. The regulator structure is not completely fixed a priori, though it can be reduced to a PID in simple cases. This is a significant peculiarity with respect to the main research stream on autotuning regulators, that refers essentially to fixed-structure (PID) regulators. Both a simulation and a laboratory example are reported.**

I. Introduction

In the industry, the term 'process controller' refers typically to a simple device, handling one or a few loops. Such devices are the most widely used in process applications (64% of the overall marketplace according to [5]), as they "provide a more manageable process in case of failure" (*ibidem*). In the selection of a process controller, autotuning is nowadays among the most desired features [6], as the users are becoming more and more confident in that technology [3, 2, 13]. Most process controllers are based on the PID regulator structure, and the diffusion of those controllers is surely among the reasons of the great research effort that in the last decades has been spent on the PID (auto)tuning [2, 13]. Methods were conceived to synthesize a PID for almost any control problem, or class of control problems, that a process controller may come across. This work addresses a control problem that is quite particular, but of significant interest in process applications, and for which it is difficult, and potentially very critical, to devise a tuning method for the PID. For this control problem, an alternative regulator structure is proposed, and a tuning procedure for that structure is devised, focusing attention on implementation-related issues. A simulation test and an experimental laboratory example are reported to illustrate the effectiveness of the proposed autotuning regulator, also by comparing it to a PID tuned with a method conceived for load disturbance rejection.

II. Problem statement

Consider a single-loop control problem with the following characteristics. The process is described by a transfer function that is stable, of type 0, essentially delay free and minimum-phase, but possibly of high order. The process may have poles and zeros that are near in frequency, and within the desired control band. The process poles are overdamped, but its step response may be overshooting. The major control goal is to counteract phenomena that are naturally modeled as unmeasured load disturbances. Not only the duration of the load disturbance response, but also the peak deviation of the controlled variable is subject to specifications.

Notice that such a problem is frequent in process control; a quite typical, but not the only example is a temperature loop with tight tolerances. In cases like this, most PID autotuning methods experience some difficulties. There is not the space here for a full discussion, but a brief sketch of the reasons of those difficulties can be given. The great majority (not to say the totality) of autotuners in process controllers use either model-based tuning methods [2, 9, 13] or relay-based identification and tuning rules that assign one point of the open-loop Nyquist curve [13, 16]. In the former case, the main problem is that simple models like those adopted in most PID tuning rules are too simple, and the role of the identification method is too critical and difficult to characterize, to acheve the bandwidths required for efficient disturbance rejection [9]. There are some remarkable exceptions, such as the "kappa-tau" (or KT) PID tuning method [2], but it is easy to verify that the regulators obtained are (necessarily) quite conservative. Things are different for relay-based autotuners, see e.g. [1, 7, 10, 12, 15, 16] and numerous other works. Most frequently, they force the open-loop Nyquist diagram to cross the unit circle with a given phase margin, that is the most natural specification in that context. The information conveyed by one point of a Nyquist diagram is local but exact, and free from hypotheses on the structure of the process dynamics. Therefore, with relay-based autotuning it is easier and safer to obtain wide control bands. However, a specification on the phase margin *per se* is not the most natural and intuitive way to express a request on load disturbance rejection, since it may have different meanings and effects in various situations.

III. The regulator

Consider the block diagram of figure 1, where $y°$ is the set point, y the controlled variable, u the control signal, d the load disturbance, $P(s)$ and $R(s)$ the transfer function of the process and the regulator, respectively.

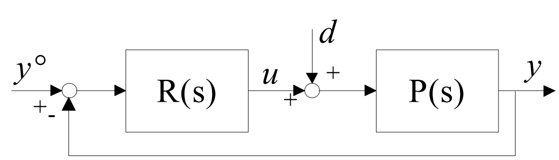

Fig. 1 The main block diagram.

0-7803-8335-4/04/$17.00 ©2004 AACC

A requirement on the rejection of d is expressed very naturally as a magnitude overbound for the frequecy response of the transfer function $G_d(s)=Y(s)/D(s)$. Since $|G_d(j\omega)|$ can be approximated with $|R(j\omega)|^{-1}$ and $|P(j\omega)|$ for frequencies ω smaller and larger than the cutoff frequency ω_c, respectively, a way to achieve good rejection of d is to design $R(s)$ so that its magnitude be as high as possible in the vicinity of ω_c. A compromise with the degree of stability is apparently in order. That compromise depends on the regulator structure, which – as a consequence – in the presented autotuner is not fixed *a priori* in its entirety. The structure is chosen so as to be able of providing a large phase lead, as its magnitude at the cutoff frequency has to be kept high [4]. Moreover, which is even more important, the structure is chosen so as to allow determining the maximum regulator lead, and the magnitude at the corresponding frequency *easily*. The structure is

$$R(s)=K\frac{(1+sT)^n}{s^m(1+s\alpha T)^{n-m}},\ m\geq 1,\ n\geq m,\ 0<\alpha<1 \quad (1)$$

Note that with m=1 and n=2, (1) is a real PID. This structure is not standard, but can be easily implemented with the elements of any control design environment. The regulator is tuned by assigning one point of the open-loop Nyquist diagram on the basis of one point of the process frequency response, identified (with a relay and/or a sine input test) at a frequency as close as possible to the desired cutoff: in [10] there is an example of such a relay-based identification method. Note, incidentally, that such a tuning policy allows to guarantee a phase margin anyway, though this is not the crucial point. Since the tuning policy is meant to use a single point of the process Nyquist curve, the number of regulator parameters is kept to the minimum necessary. The regulator (1) gives the maximum phase lead at the frequency

$$\overline{\omega}(\alpha,m,n,T)=\frac{1}{T}\chi(\alpha,m,n), \quad (2)$$

where

$$\chi(\alpha,m,n)=\sqrt{\frac{\alpha(n-m)-n}{\alpha(\alpha n+m-n)}}, \quad (3)$$

under the condition $0<\alpha<(n-m)/n$, and that maximum phase lead is given by

$$\begin{aligned}\overline{\varphi}(\alpha,m,n) &= n\,arc\,tan(\chi(\alpha,m,n))-m\,90°+\\ &-(n-m)arc\,tan(\alpha\chi(\alpha,m,n)).\end{aligned} \quad (4)$$

The regulator magnitude at the frequency of the maximum phase lead is

$$\overline{R}(\alpha,m,n,T,K)=\frac{KT^m\left(1+\chi(\alpha,m,n)^2\right)^{\frac{n}{2}}}{\chi(\alpha,m,n)^m\left(1+(\alpha\chi(\alpha,m,n))^2\right)^{\frac{n-m}{2}}}, \quad (5)$$

while the asymptotic regulator magnitude for $\omega\to\infty$ is

$$R_\infty(\alpha,T,K)=\frac{KT^m}{\alpha^{n-m}}. \quad (6)$$

The typical aspect of the magnitude and phase of $R(j\omega)$ is shown in figure 2 (obtained with m=1 and n=3). For better clarity, since the figure is essentially qualitative, the frequency axis is normalized - i.e., graduated in ωT - and the magnitude plot is scaled as if KT^m=1.

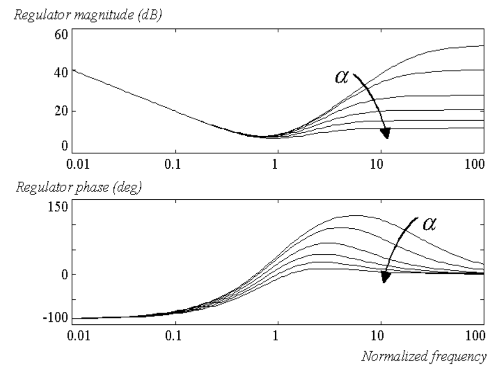

Fig. 2 Regulator's magnitude and phase.

IV. A TUNING PROCEDURE

The idea behind the tuning is very simple. Suppose that one point $P(j\omega_o)=A_P e^{j\varphi_P}$ of the process Nyquist curve has been identified with a relay experiment, ω_o being the limit cycle frequency, that a phase margin φ_m is required, and that the regulator magnitude at the cutoff frequency must be at least R_{min}. A regulator R(s) in the form (1) solves the problem, making ω_o also the cutoff frequency, if

$$\begin{cases}Ae^{j\varphi_P}R(j\omega_o)=e^{j(\varphi_m-180°)}\\ |R(j\omega_o)|\geq R_{min}\end{cases} \quad (7)$$

On the basis of the definitions above, this is possible if

$$\begin{cases}\varphi_m-180°-\overline{\varphi}(\alpha,m,n)\leq\varphi_P\leq\varphi_m-m90°\\ A_P\leq 1/R_{min}\end{cases} \quad (8)$$

In other words, (8) define the locus Π of the points of the process Nyquist curve that the regulator (1) is capable of moving to $e^{j(\varphi m-180°)}$ (obtaining φ_m) but, and this is the key point of the proposed method, with a sufficiently large magnitude of R(jω) at ω_o; this is illustrated in figure 3, where Π is evidenced by the darker hatch.

The goal of any tuning procedure based on the proposed approach is quite articulated. In fact, it is necessary to find a point of the process Nyquist curve such thatits frequency is as close as possible to the desired ω_o,its magnitude is less than 1/ R_{min}, given a value of m (based on *a priori* static specifications) there exist α and n such that the point is contained in Π, and if the regulator is synthesised by moving that point to $e^{j(\varphi m-180°)}$, i.e., determining K and T with the first (complex) equation of (7), the resulting regulator's high frequency gain R_∞ is not too high. Notice that an upper bound on R_∞ reflects in one on n.

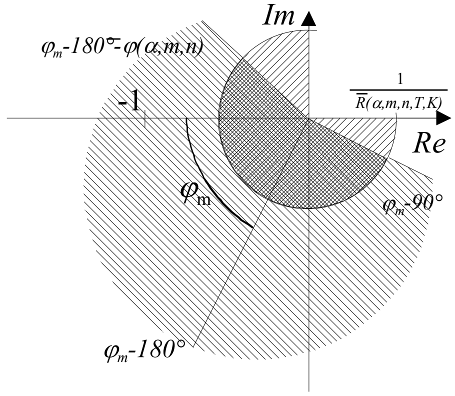

Fig. 3 Region of 'suitable' points of *P(jω)*.

Apparently, this is a multi-objective problem, and in real-life cases it may be impossible to solve, either because specifications are too strict or because the synthesis approach is not suited to the problem. Therefore, the tuning procedure must be capable of relaxing specifications if required, and provide an alternative tuning if its approach is inherently unfit for the problem at hand. In both these cases, a warning to the user is necessary. Such a procedure can be designed in many different ways, and in any case it turns out to be complex. For this reason, instead of giving a lengthy description of the procedure adopted to generate the presented results, in the following some guidelines are given to design a possible tuning procedure.

First, suppose that no cutoff frequency is specified. In this case, find the process ultimate point, which means setting ω_o to the ultimate frequency. If $|P(j\omega_o)|<1/R_{min}$, choose n so that there surely exists a value of α for which $P(j\omega_o)$ is contained in region Π (see figure 3). In other words, since the required regulator magnitude at the cutoff can be obtained, choose the regulator structure so as to be able of providing the necessary phase lead. To choose n, recall that a higher value will result in a larger α, hence in less distance between the poles and zeros of R(s). In addition, n influences R_∞ through (6), but in this relationship also T is present. Therefore, a reasonable choice is to first determine the required phase lead at ω_o (in any reasonable application, a lead is necessary) based on $P(j\omega_o)$ and the required phase margin φ_m, and then introduce one zero for every 40° of required lead. Notice that, if the ultimate point magnitude is acceptable, the regulator will be a real PID with coincident zeros if the required lead (i.e., φ_m) is less than 40°, and for $\varphi_m<80°$ n will not exceed 3. Introducing one regulator zero for every 40° of phase lead is a good policy to avoid high values of α. Once n is computed, set α so that the maximum regulator lead equals the required lead (plus a few degrees for numerical reasons), i.e., so that $P(j\omega_o)$ be inside Π but near its boundary; then, set T so that this lead occur at ω_o. This is accomplished based on (2) and (3). In so doing, sensitivity will be kept low in the vicinity of the cutoff frequency, which is a highly desirable characteristic of any control loop, and this will be achieved without an excessive high frequency regulator gain (see figure 2). It must be observed that this way of determining the maximum regulator lead and α makes the proposed method not very suited for systems with dominant delay, however. Once α and T are available, Compute K so that $|R(j\omega_o)P(j\omega_o)|=1$. If the resulting R_∞ is not too large (a bound is determined easily based on the measurement noise), the tuning is achieved. In any other case, the ultimate point is not 'suitable'. This may mean that its magnitude is not small enough and/or that R_∞ is too large.

If, the process ultimate magnitude is not small enough, try increasing ω_o, identify the corresponding point, and recompute in sequence n, α, T, K, and R_∞. If R_∞ is not too high, verify that the process magnitude is less than $1/R_{min}$: if this is true the tuning is achieved. In the opposite case, record the quantities computed, and particularly the achievable value of R_{min} (which is $1/|P(j\omega_o)|$, and was not acceptable), then increase the cutoff again, and repeat the identification, the computations and the checks until a value of 3 times the ultimate frequency is reached. At this step, if a suitable point has not been found yet, there are the following possibilities.

a) The ultimate magnitude was too high and no point yielded a sufficiently small R_∞ (i.e., n) This is likely to indicate a dominant delay, as the process phase appears to diminish rapidly and steadily above the ultimate frequency. In this case, the proposed approach is not suited to the process: however, one could tune with the ultimate point, and alert the user that the required R_{min} was not attained. Results will be similar to standard one-point relay based tuning.
b) The ultimate magnitude was too high, some points gave a sufficiently small R_∞, but the process magnitudes were too high. This indicates that the proposed approach is suited for the process, but the specification on R_{min} was too demanding. In this case, one could tune with the point yielding the largest R_{min} (values were recorded for this purpose) and, again, warn the user.

If the problem with the ultimate point was not the magnitude, but the excessive lead required, it is possible to reduce ω_o. The regulator parameters are computed in the same way, and the procedure is equivalent to standard one-point PID tuning, as witnessed by example 3. If a desired cutoff frequency is specified, the procedure is the same, but the first point is identified at the desired cutoff frequency, e.g. with the method of [10] and not at the ultimate frequency. Notice that most frequency search techniques also provide at least a guess of the ultimate point.

It is worth stressing that the procedure sketched above does not exhaust the possibilities of the approach; one could devise more complex procedures, for example observing how the process magnitude and phase vary when points at increasing frequencies are identified and drawing some conclusions on the presence and entity of a possible delay, or allowing the procedure to modify also the phase margin, and so on. Discussing all these possibilities would be lengthy and inessential, however. The sketch above was presented to show that the proposed approach is applicable, and heuristics (i.e., default values and thresholds) can be given a sense. The interested reader is encouraged to experiment with the approach, and devise different procedures.

V. A SIMULATION EXAMPLE

The process considered in this example is described by the transfer function

$$P(s) = \frac{1}{(1+s)^4}.$$

Employing the proposed method with the ultimate frequency (the point found has ω_o=1.733, A_P =0.125, φ_P =-180.5°), with φ_m=45° and setting m=1, gives n=2, α=0.082, T=3.047, K=0.524; the actual phase margin obtained with this regulator is 45.008°. With φ_m=30° the result is n=2, α=0.155, T=2.394, K=0.907; the actual phase margin is 30.023°. Notice that these two regulators are real PIDs.

Then, the method was applied identifying the point at 3 times the ultimate frequency. It would be extremely difficult to do this with a relay, therefore a sinusoidal input was used. The point found in this way has ω_o=1.733, A_P=0.007, φ_P=-237.4°. With φ_m=60° and m=1 the result is n=3, α=0.051, T=1.064, K=4.717. With φ_m=45°, the method yields n=3, α=0.080, T=0.862, K=9.050; the actual phase margin in the two cases is 60.112° and 45.184°.

To compare the proposed method with a well established and effective one, widely employed in the application domain and particularly suited for load disturbance rejection problems, the KT tuning rules [2], that refer to the 1-d.o.f. ISA PID control law

$$R_{PID}(s) = K_P\left[1 + \frac{1}{sT_i} + \frac{sT_d}{1 + sT_d / N}\right],$$

were applied to a model of First-Order Plus Dead Time (FOPDT), i.e., in the form $M(s)=\mu_M e^{-sL_M}/(1+sT_M)$, identified with the well known method of areas, that gave μ_M=1.0, T_M=1.85, and L_M=1.15. Setting the design parameter M_s of the KT method (that is related to the maximum sensitivity, see [2] for a complete explanation of the method) to 1.4, the KT method gives K_P=0.707, T_i=2.093, and T_d=0.486 (ω_c=0.327, φ_m=73.34°), while setting M_s=2.0 gives K_P=1.426, T_i=2.028, T_d=0.448 (ω_c=0.619 and φ_m=57.591°). Parameter N was set to 5 in both regulators (the value of N is of neglectable importance for the problem addressed). The closed-loop load disturbance responses of the various control systems are shown in figure 4 (notice the different amplitude and time scales of the various plots). Strictly speaking, only the comparison between the first two couples of plots is meaningful, since the KT rules tune a PID. Even in this case, however, the results obtained with the proposed technique are better, especially considering the peak deviation of the controlled variable, and not only the transient's duration. Things are very different when the PID structure is abandoned (third couple of plots). In that case, the advantages of adopting the proposed regulator structure and the corresponding tuning technique are apparent.

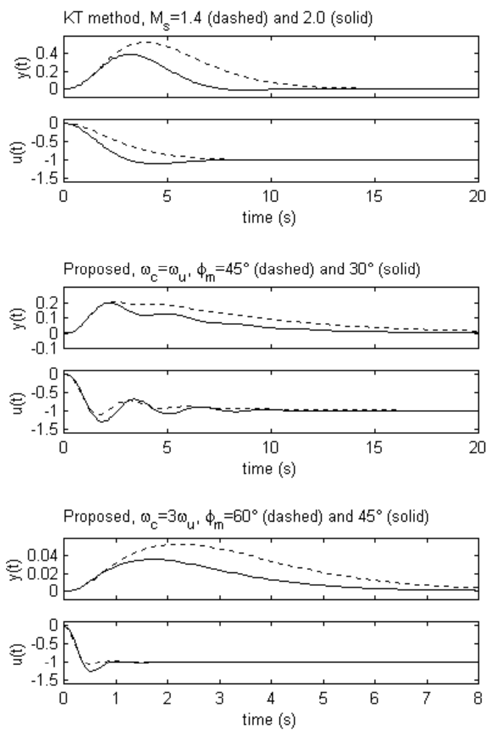

Fig. 4. Responses in the simulation example.

VI. A LABORATORY APPLICATION

In this section, the proposed autotuner is applied to a laboratory apparatus. The apparatus, described in [11], is composed of a small metal plate, heated by two transistors and cooled by a fan. The inputs are the two transistor commands (in percent of the maximum thermal power) and the on-off fan command, that in the presented experiment is not used. The outputs are the temperatures of the two transistors and the metal plate, in °C. In the experiment presented, the controlled variable is the plate temperature, one transistor is used as the control actuator, and the other transistor provides a load disturbance. The experiment begins with a relay test, that determines the ultimate point (ω_o=0.36, A_P=0.0021). Then, a sine input test at 1.5 times the ultimate frequency gives ω_o=0.54, A_P=0.001, φ_P=-195.13°. Selecting 1.5 times the ultimate frequency is a reasonable choice when the signals are noisy, as in this case.

The identification experiment is shown in figure 5. The relay amplitude was set *a priori* (as is common practice in relay-based autotuning process controllers), while the amplitude of the subsequent sine wave was chosen so as to yield an output oscillation of sufficient amplitude, assuming that the process has essentially a low-pass behaviour in the band of interest (which is reasonable).

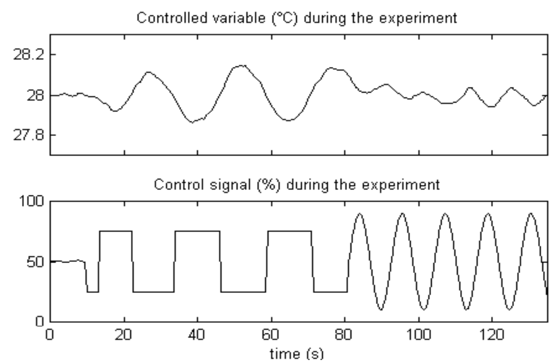

Fig. 5. Identification experiment.

Notice that choosing the switching time between the relay and the sine wave in a convenient way (i.e., at a phase lag of 45° with respect to the last peak of the oscillation induced by the relay, according to the period estimated with the previous oscillation) results in an acceptable experiment length. The case presented is particularly good, but no experiment ever required more than 8 periods of the sine input to allow the identification of the second point. Further details on the realization of the experiment are omitted for brevity. With φ_m=45° and *m*=1 the proposed method yields *n*=2, α=0.035, *T*=14.4, *K*=10.26; with φ_m=30° the result is *n*=2, α=0.082, *T*= 9.81, *K*=22.8. These two regulators are compared in figure 6 with two PIDs tuned with the KT method and a FOPDT model $M_1(s)$, having μ_P=0.18 T_P=102.8, L_P=7.5. The transient considered is the response to a 50° load disturbance step.

To appreciate the criticality of identification in model-based methods, the KT rules were applied to two other models, termed $M_2(s)$ and $M_3(s)$, one with μ_P=0.18, T_P=98.6, L_P=12.5, and the other with μ_P=0.18, T_P=125.6, L_P=8.8. These models are within the variability of realistic experiments, and they all fit the measured data.

Using the KT method with $M_2(s)$ and $M_3(s)$ produces the experimental transients shown in figure 7. Numbers are omitted for brevity, but it is immediate to see that in a case like this the influence of the identification can be very critical. The proposed method, that guarantees a phase margin anyway, is less sensitive to identification errors.

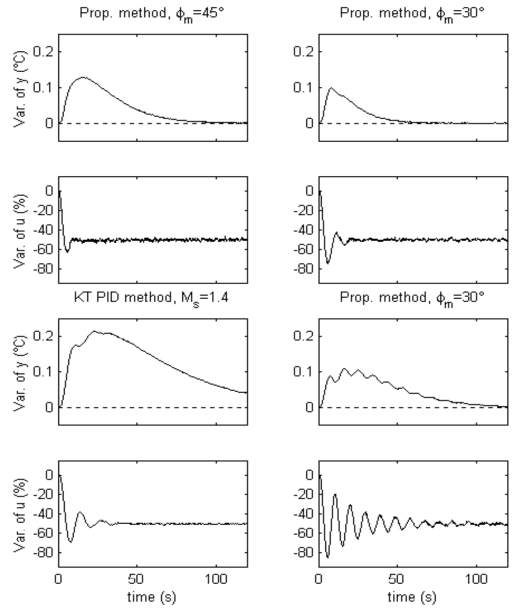

Fig. 6. Experimental tuning results.

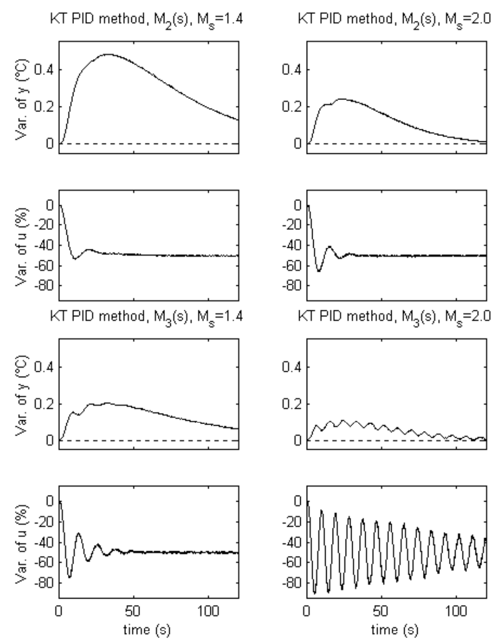

Fig. 7. Results of the KT method with different models.

It would be interesting to discuss also the comparison between the presented autotuner and standard relay-based PID autotuners. This would exceed space limitations, however, as there are many different methods that should be addressed. Therefore, only some remarks will be given here, to sketch out the reason why, for the load disturbance rejection problem addressed, the presented autotuner is preferable to standard relay-based PID ones. Basically, the reason is that the typical relay-based PID autotuners normally operate at frequencies lower than the ultimate one, having as principal objective the phase margin. In some cases, the point of the process Nyquist curve with phase –90° is used, for example: this may provide a better stability and robustness degree than the presented method, but at the cost of a correspondingly increased conservatism, and of less performance. From the standpoint of disturbance rejection, the presented autotuner provides an improvement in practically the totality of the cases addressed.

VII Some words on implementation

The proposed method is quite simple and light from the computational point of view. In the implementation it is necessary to take the usual precautions when doing relay or sine input experiments, but nothing is required in this respect that is not explained in the vast literature on those subjects.

However, any autotuner has to provide a reasonable default for its design parameters. Here, apart from m that is the unity in any realistic case, the only design parameters are the phase margin (a good default being 45°, as already shown, and a sensible 'aggressive' value being 30°) and the desired cutoff frequency. The bound on R_∞ can be selected automatically, based on possible limitations on the variation of the control signal and on the measured noise amplitude: also this matter is extensively discussed in the PID autotuning literature, and from this point of view the proposed structure is analogous to a rela PID.

Quite intuitively, it is advisable (and easy to understand for the operators) to relate the desired ω_c to the ultimate frequency: 1.5 times the ultimate frequency can be taken as a reasonable default, and three times that frequency can be considered the maximum allowed with very small measurement noises. Notice also that the ultimate frequency is easy and quick to determine with a single relay test, so that finding a second point (typically with a sine test) preserves a short duration of the overall experiment.

VIII. Conclusions

An autotuning process controller specifically aimed at rejecting load disturbances was presented. The class of control problems considered is relevant in process control and, as shown by the comparative examples reported, it can be critical for the majority of industrial autotuners, that refer to the PID controller structure.

The structure adopted in this work is slightly different from the PID, although in simple cases it can be reduced to a PID (the tuning method does this automatically, as shown). It is a simple structure, however, with a very small number of parameters, and easy to implement in any industrial control environment with standard elements.

The regulator synthesis is done by assigning one point of the open-loop Nyquist diagram on the basis of one point of the process frequency response (identified by means of a conveniently driven relay experiment and/or a sine input test), keeping the regulator magnitude at the cutoff frequency as large as possible.

In the control problems it was devised for, the autotuner can yield significant improvements, both in terms of performance and criticality of the identification phase, with respect to the available alternatives, as demonstrated by the tests reported.

IX. References

[1] K.J. Åström and T. Hägglund, "Automatic Tuning of Simple Regulators with Specifications on Phase and Amplitude Margins", *Automatica*, vol. 20, No. 5, 1984, pp. 645-651

[2] K.J. Åström and T. Hägglund, *PID Controllers: Theory, Design and Tuning - Second Edition*, Instrument Society of America, Research Triangle park, NY; 1995.

[3] K.J. Åström, C.C. Hang, P. Persson and W.K. Ho, "Towards Intelligent PID Control", *Automatica*, vol. 28, No. 1, 1992, pp. 1-9.

[4] S.P. Boyd and C.H. Barratt, *Linear Controller Design: Limits of Performance*, Prentice-Hall, 1991.

[5] Control Engineering, May 1998, *Single-loop Controllers Dominate Marketplace*.

[6] EnTech, 1994, *Competency in Process Control – Industry Guidelines, version 1.0*.

[7] W.K. Ho, C.C. Hang and L.S. Cao, "Tuning of PID Controllers based on Gain and Phase Margin Specifications", *Automatica*, vol. 31, No. 3, 1995, pp. 497-502.

[8] A. Leva, "How to Construct PID Synthesis Rules for Structurally Complex Process Models", *in SSC '98*, Nantes, F, 1998.

[9] A. Leva, "Model-based Tuning: the Very Basics and some Useful Techniques", *Journal-A*, vol. 42, No. 3, 2001, pp. 14-22.

[10] A. Leva, "PID Autotuning Algorithm based on Relay Feedback", *IEE Proceedings-D*, vol. 140, No. 5, 1993, pp. 328-338.

[11] A. Leva, "A Hands-on Experimental Laboratory for Undergraduate Courses in Automatic Control", *IEEE Transactions on Education*, vol. 46, No. 2, 2003, pp. 263-272.

[12] W.L. Luyben, "Getting More Information from Relay Feedback Tests", *Ind. Eng. Chem. Res.*, vol. 40, No. 20, 2001, pp. 4391-4402.

[13] A. O'Dwyer, *Handbook of PI and PID Controller Tuning Rules*, World Scientific Publishing; 2003.

[14] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control - Analysis and Design*, Wiley, 1986.

[15] T. Thyagarajan and C.C. Yu, "Improved Autotuning Using the Shape Factor from Relay Feedback", *Ind. Eng. Chem. Res.*, 2002, DOI 10.1021/ie011006f.

[16] C.C. Yu, *Autotuning of PID Controllers: Relay Feedback Approach*, Springer-Verlag, London, UK; 1999.