Preprints, 4th IFAC Conference on
Advances in Proportional-Integral-Derivative Control
Almería, Spain | June 12-14, 2024

Robust PID controller tuning using bilevel optimization

Lucas Cammann, Johannes Jäschke

Department of Chemical Engineering, Norwegian University of Science and Technology Trondheim, NO-7491 (e-mail: lucas.cammann@ntnu.no, johannes.jaeschke@ntnu.no).

Abstract: Tuning PID controllers for performance while satisfying robustness constraints is a difficult optimization problem. Previous authors have resorted to fine frequency discretization or global optimization algorithms to ensure robustness constraint satisfaction. In this work, we propose to formulate the tuning problem as a bilevel optimization problem, where the upper level cost function quantifies control performance, and the lower level problem defines frequencies at which robustness must be ensured. Here, the first-order optimality conditions for the peaks of the sensitivity functions (the lower level) are embedded as constraints to the performance optimization problem. Automatic differentiation is used to construct the problem and for efficient numerical solving, with convergence typically achieved in the order of 0.1 seconds. The methodology is applied to two case studies, where it is found to be suitable to improve existing controller designs both in terms of performance and robustness.

Keywords: PID tuning, Bilevel optimization, Time delay, Robustness constraints

1. INTRODUCTION

Proportional-Integral-Derivative (PID) controllers are abundantly used in industry. Although they are classically defined with only three parameters, finding good tunings is a non-trivial task. Over the years, several tuning methods have been proposed. Examples include the (rather poorly performing, but famous) Ziegler-Nichols rules (Ziegler and Nichols, 1942), or the more recent SIMC-rules (Skogestad, 2003). While the latter offer good solutions for a defined subset of problems, more complex tuning problems warrant the use of optimization-based methods. This is especially the case when considerations on the trade-off between performance and robustness are of importance, or problem classes in which the tuning rules are not applicable.

An early work describing an optimization approach for this trade-off is Balchen (1958), where the integrated absolute error (IAE) is minimized subject to constraints on peaks of the sensitivity function. This trade-off has seen continued interest since, with Garpinger et al. (2014) and Grimholt and Skogestad (2018a) calculating trade-off curves for different PID-controlled systems. To improve convergence, the latter derived exact gradients for the objective and constraint functions for a certain class of systems and enforced the robustness constraints on a finely discretized frequency grid (Grimholt and Skogestad, 2018b). To avoid the computational cost of this discretization, Turan et al. (2022) used a global optimization algorithm showing similar results in the same test case.

The aforementioned works show reliable performance but are also computationally expensive and require upfront work by the designer to be adapted. In this work we propose a simplified approach, in which the PID tuning problem is locally solved using a bilevel optimization formulation. Specifically, we reformulate the bilevel problem into one single level by formulating the optimality conditions of the sensitivity function peaks as constraints. Importantly, the developed method makes use of automatically generated symbolic derivatives allowing for adaptable problem formulation and efficient numerical solving. We showcase the method in two case studies chosen to be particularly challenging for this kind of approach. Here it is found to be successful in improving existing designs in terms of performance, or in terms of robustness with rapid solve times.

The remainder of this work is structured as follows. First, the theoretical background at the basis of this work is introduced. Afterwards, the generic bilevel problem formulation used for the PID tuning is presented. Following that, the algorithm employed in this work is detailed before showcasing numerical results from two case studies. In doing so, benefits and limitations of the proposed method are discussed alongside the presented results. Finally, concluding remarks are made.

2. THEORETICAL BACKGROUND

The following section introduces the theoretical background relevant for the exposition of this work. The generic feedback control scheme is introduced together with the definition of relevant sensitivity functions, before the PID controller is introduced. Lastly, concepts in bilevel optimization used in this work are briefly described.

2.1 Feedback control

Fig. 1 shows the block diagram of a one-degree of freedom feedback control system. Here, K is the controller and G the plant transfer function. The signals d_{y} and d_{u} are

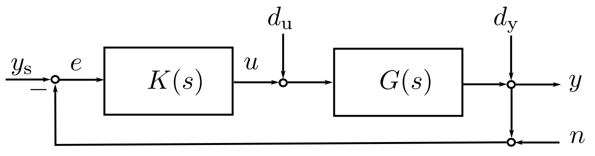


Fig. 1. Considered block diagram of the closed loop system.

the disturbances entering at the plant output and input respectively, while n is the measurement noise. The setpoint y_s is assumed zero going forward. Upon rearranging the equations resulting from the block diagram structure, the effects of these signals on the control error e as well as the plant input u may then be written as

$$-e = y - y_s = S(s)d_y + GS(s)d_u - T(s)n, \quad (1)$$
$$-u = KS(s)d_y + T(s)d_u + KS(s)n. \quad (2)$$

The transfer functions multiplying the disturbance and noise signals in Eqs. 1 and 2 are generally defined as follows for a multivariable system

$$S(s) = (1 + G(s)K(s))^{-1}, \qquad T(s) = 1 - S(s), \quad (3)$$
$$GS(s) = G(s)S(s), \qquad KS(s) = K(s)S(s).$$

The transfer function S is known as the sensitivity function, and T as the complementary sensitivity function. Above, the sensitivity function S can be seen as the closed-loop transfer-function from the output disturbance to the output, and T analogously from the reference signal to the outputs. Hence, good disturbance rejection and command tracking are achieved with $|S| \approx 0$, and $|T| \approx 1$ respectively. On the other hand, achieving low noise transmission requires $|T| \approx 0$ and $|S| \approx 1$, which stands in contrast to the previous requirements. Furthermore, it is desirable to keep the magnitudes $|KS|$ small to reduce the control signal.

2.2 Robustness and performance

In designing feedback controllers, restrictions on the robustness of the closed-loop system are commonly defined in terms of the aforementioned sensitivity functions. For the examples covered in this work, these robustness requirement are based on the largest sensitivity peak $M_{ST} = \max(M_S, M_T)$ (Grimholt and Skogestad, 2018b), with

$$M_S = \max_{\omega} |S(j\omega)| = \|S(j\omega)\|_\infty \quad (4)$$
$$M_T = \max_{\omega} |T(j\omega)| = \|T(j\omega)\|_\infty.$$

Here $\|\cdot\|_\infty$ is defined as the $\mathcal{H}_\infty$ norm, i.e. the maximal magnitude as a function of frequencey ω. Generally, small values of M_{ST} are desired for robustness with Grimholt and Skogestad (2018b) recommending M_S be no larger than 2. Improving the robustness of the closed-loop system usually comes at the expense of performance in terms of disturbance rejection (Åström and Hägglund, 2006). Most commonly, the performance is expressed in terms of the integral absolute error (IAE)

$$IAE_d = \int_0^\infty |e_d(t)| dt. \quad (5)$$

Here, the subscript d might refer to any disturbance or change in reference. As such, Grimholt and Skogestad (2018b) consider an objective function combining the IAE of both the input and output disturbances

$$J = 0.5\big(\Phi_{dy} IAE_{dy} + \Phi_{du} IAE_{du}\big), \quad (6)$$

where Φ_{dy} and Φ_{du} are normalization factors. Using the IAE of Eq. 5 as performance metric has the downside of requiring an expression of the error function in the time domain. In this work, an approximation of the IAE in the frequency domain is used, which goes back to Balchen (1958)

$$IAE_d = \int_0^\infty |e_d(t)| dt \approx \max_{a,w} \int_0^\infty e_d(t) \sin(wt + a) dt \quad (7)$$
$$\approx \|E_d(s)\|_\infty = HIE_d. \quad (8)$$

Here, the term HIE_d denotes the $\mathcal{H}_\infty$ error of the disturbance signal d, e.g. it is equivalent to $\|S(s)\|_\infty$ for a output disturbance d_y. The HIE is a lower approximation of the IAE ($HIE \leq IAE$), and was found to give quite good approximations for well-dampened systems.

2.3 PID control

The by far most commonly employed controller in industrial-practice is the PID controller. In the parallel form, its transfer function is defined as

$$K(s) = K_p + K_i/s + K_d s, \quad (9)$$

with K_p, K_i and K_d being the proportional, integral and derivative gain, respectively. Going forward, the notation of Grimholt and Skogestad (2018b) is adapted with the controller parametrization being

$$p = [K_p,\ K_i,\ K_d]. \quad (10)$$

2.4 Bilevel optimization

Bilevel optimization problems are optimization problems which are characterized by containing another optimization problem as a constraint. Following the nomenclature of Sinha et al. (2018), such problems may generally be written as

$$\begin{aligned} &\min_{x_u \in X_U, x_l \in X_L} F(x_u, x_l) && (11) \\ &\text{s.t.} \quad x_l \in \operatorname*{argmin}_{x_l \in X_L} \left\{ f(x_u, x_l) : g_j^l(x_u, x_l) \leq 0, j = 1, \ldots, J \right\} \\ &\qquad g_k^u(x_u, x_l) \leq 0, k = 1, \ldots, K \end{aligned}$$

where x_u denote the upper, and x_l the lower level decision variables respectively and g_k^u and g_j^l are upper and lower level constraints. Substituting the lower problem by its first-order optimality conditions, the single level reduction of Eq. 11 is obtained as

$$\begin{aligned} &\min_{x_u \in X_U, x_l \in X_L, \lambda} F(x_u, x_l) && (12) \\ &\text{s.t.} \quad \nabla_{x_l} \mathcal{L}(x_u, x_l, \lambda) && = 0 \\ &\qquad g_k^u(x_u, x_l) && \leq 0,\ k = 1, \ldots, K \\ &\qquad g_j^l(x_u, x_l) && \leq 0,\ j = 1, \ldots, J \\ &\qquad \lambda_j g_j^l(x_u, x_l) && = 0,\ j = 1, \ldots, J \\ &\qquad \lambda_j && \geq 0,\ j = 1, \ldots, J. \end{aligned}$$

Above, the lower level problem has been replaced by its respective Karush-Kuhn-Tucker (KKT) conditions. This entails the gradient of the Lagrangian $\mathcal{L} = f(x_u, x_l) + \sum_j g_j^l(x_u, x_l)\lambda_j$, as well as complementarity constraints (CC) of the form $\lambda_j g_j^l = 0$, and dual feasibility (DF) of the form $\lambda_j \geq 0$. Here, λ denotes the Lagrange multiplier. Generally, problems of this form are not easy to solve due to the combinatorial nature of the complementarity constraints. Reformulation strategies are hence required when using nonlinear program (NLP) solvers. One such reformulation is the penalty function method, in which the complementary slackness is appended to the objective of the upper level problem as a weighted penalty.

3. BILEVEL OPTIMIZATION FOR PID TUNING

The following section presents the mathematical formulation of the controller tuning procedure proposed in this work.

3.1 General problem formulation

The intuition behind the proposed approach stems from the realization that the infinity norms of the considered sensitivity functions are themselves defined by an optimization problem and can hence be casted into a suitable bilevel optimization framework. In this work, it is sought to minimize the performance index of the HIE, while satisfying certain robustness constraints. Put concisely into the form of Eq. 11 this amounts to

$$\min_{\omega_\mathrm{l},\lambda,p} \sum_{\mathrm{d}=1}^{\mathrm{D}} \Phi_\mathrm{d} HIE_\mathrm{d}(\omega_\mathrm{E}, p) \tag{13}$$

$$\text{s.t.} \left.\begin{aligned} &\omega_\mathrm{S} \in \operatorname{argmax}\{|S| : \omega_\mathrm{lb} \le \omega_\mathrm{S} \le \omega_\mathrm{ub}\} \\ &\omega_\mathrm{T} \in \operatorname{argmax}\{|T| : \omega_\mathrm{lb} \le \omega_\mathrm{T} \le \omega_\mathrm{ub}\} \\ &\omega_\mathrm{E} \in \operatorname{argmax}\{|E_\mathrm{d}| : \omega_\mathrm{lb} \le \omega_\mathrm{E} \le \omega_\mathrm{ub}\} \end{aligned}\right\} \text{Lower level peaks}$$

$$\left.\begin{aligned} &p_\mathrm{lb} \le p \le p_\mathrm{ub} \\ &M_\mathrm{ST} - M_\mathrm{ub} \le 0 \end{aligned}\right\} \text{Upper level inequalities } g^\mathrm{u}.$$

Here, D is the number of HIE terms considered in the objective, e.g. two in the case of weighted input and output disturbances. The subscripts lb and ub refer to the lower and upper bounds. The upper level decision variables are the controller parameters p, with the upper level constraint set g^u consisting of bounds on M_ST and on p. Applying the single level reduction to Eq. 13 gives

$$\min_{\omega_\mathrm{l},\lambda_\mathrm{l},\mathrm{p}} \sum_{\mathrm{d}=1}^{\mathrm{D}} \Phi_\mathrm{d} HIE_\mathrm{d}(\omega_\mathrm{E}, p) \tag{14}$$

$$\text{s.t.} \left.\begin{aligned} &\frac{\partial|S(\omega,p)|}{\partial\omega}\Big|_{\omega_\mathrm{S}} + \lambda_1 - \lambda_2 &= 0 \\ &\frac{\partial|T(\omega,p)|}{\partial\omega}\Big|_{\omega_\mathrm{T}} + \lambda_3 - \lambda_4 &= 0 \\ &\frac{\partial|E_\mathrm{d}(\omega,p)|}{\partial\omega}\Big|_{\omega_\mathrm{E}} + \lambda_\mathrm{d} - \lambda_\mathrm{d+2} &= 0 \end{aligned}\right\} \nabla_{\omega_\mathrm{l}}\mathcal{L}$$

$$\left.\begin{aligned} &|S(\omega_\mathrm{S}, p)| - M_\mathrm{ub} &\le 0 \\ &|T(\omega_\mathrm{T}, p)| - M_\mathrm{ub} &\le 0 \\ &p_\mathrm{lb} \le p \le p_\mathrm{ub} \end{aligned}\right\} g^\mathrm{u}$$

$$\omega_\mathrm{lb} \le \omega_\mathrm{l} \le \omega_\mathrm{ub}\, \forall\, l \in [S, T, E]\} \quad g^\mathrm{l}$$

$$\lambda_\mathrm{j} g^\mathrm{l}_\mathrm{j}(\omega_\mathrm{l}) = 0,\ \lambda_\mathrm{j} \ge 0,\ j = 1, \ldots, J\} \quad \text{CC, DF.}$$

The lower level decision variables are gathered as $\omega_\mathrm{l} = [\omega_\mathrm{S}, \omega_\mathrm{T}, \omega_\mathrm{E}]$, of which the respective bounds are formulated in the lower level constraint set g_l. Note here that each sensitivity function peak is defined by its own Lagrangian, containing only multipliers for the respective upper and lower frequency bounds. Going forward, any active constraints on the frequency range in ω_l are assumed to be practically weakly active ($\lambda_\mathrm{i} \approx 0$). This means that the lower level problem is assumed to be unconstrained and that the complementarity constraints and dual feasibility are omitted from Eq. 14. It can be intuitively understood that extremal values of sensitivity functions at properly chosen boundaries assume asymptotic values for strictly proper systems. For example, $S = 1$ as $\omega \to \infty$ and will be close enough to 1 at high frequencies to not influence the objective function.

3.2 Constraint qualifications

For a KKT-point x^* found solving Eq. 14 to represent a necessary optimality condition, a constraint qualification such as linear independence constraint qualification (LICQ) must hold. LICQ is satisfied when the gradients of the active inequality constraints, and the gradients of the equality constraints are linearly independent. Below, an outline of argumentation is presented that can be used to easily verify the linear independence constraint qualification. For the sake of the presentation, we consider the controller parameters and the frequencies to be unconstrained and an objective of HIE_dy. The only equality constraints are then the stationarity conditions, with the inequality constraints being sensitivity magnitudes in g_u. In the most improbable case, all inequality constraints are equally active, giving the following Jacobians with respect to $x = [\omega_\mathrm{P}, \omega_\mathrm{T}, p]$

$$\nabla g_\mathrm{u}(x^*) = \begin{bmatrix} 0 & 0 & \frac{\partial|S|}{\partial p} \\ 0 & 0 & \frac{\partial|T|}{\partial p} \end{bmatrix} \nabla^2\mathcal{L}(x^*) = \begin{bmatrix} \frac{\partial^2|S|}{\partial\omega_\mathrm{P}^2} & 0 & \frac{\partial^2|S|}{\partial\omega_\mathrm{P}\partial p} \\ 0 & \frac{\partial^2|T|}{\partial\omega_\mathrm{T}^2} & \frac{\partial^2|T|}{\partial\omega_\mathrm{T}\partial p} \end{bmatrix} \tag{15}$$

In Eq. 15 the fact that the derivative of the sensitivity functions vanishes at the critical point has been used to obtain a simplified expression. Upon inspection of Eq. 15, it can be seen that linear constraint independence can be assessed based on readily available derivative information. That is, LICQ holds if

(1) The derivatives of the sensitivity functions with respect to the parameters are not multiples of each other
(2) The second derivatives of the sensitivity functions with respect to the frequencies are non-zero.

Similar arguments can be constructed with more sensitivity functions and constraints and checked with available symbolic derivatives. It has to be noted that the satisfaction of LICQ for this problem is in general conditional on the omission of the complementarity constraints, as argued above.

3.3 Convexity and uniqueness of solution

The problem posed in Eq. 14 is generally non-convex, which has several consequences. For one, it implies that different initial guesses may lead to solutions of varying quality. Hence, care must be taken in choosing appropriate initial guesses for the optimization variables. Furthermore, the numerical algorithm might find an extremal to the lower level problem which itself does not constitute the global maximum of the sensitivity functions. This circumstance might arise in case of repeated peaks, which has motivated Grimholt and Skogestad (2018b) to apply a finely discretized frequency grid and Turan et al. (2022) to apply a global optimization algorithm. It is thus imperative to check the solution returned by the solver for feasibility, and a robust algorithm based on Eq. 14 then includes such checks and other regularizations. Some of these are discussed in Sec. 4. While this renders the implementation more involved, it is not seen as a downside compared to aforementioned approaches which are more complex by design. In this sense, our method favors simplicity over strong guarantees and corrects afterwards for potential infeasibilities when they are detected.

4. NUMERICAL IMPLEMENTATION

The following section presents the numerical implementation of the proposed tuning approach. To this end, the general algorithm is outlined before further implementation and software details are presented.

4.1 Algorithm

The following algorithm is implemented in this work. Starting from an initial guess, a plant transfer function G and desired tolerances `tol`, the derivatives of the transfer functions are found symbolically, as well as the Jacobians and Hessians of the constraint and objective functions. The symbolic differentiation is a decisive feature of this algorithm as it gives analytic expressions for the lower level optimality conditions, allows for more efficient numerical solving and rapid verification of the constraint qualification. Since the optimization variables are embedded symbolically in the mathematical representation, the initial guesses can be adapted after the problem formulation has been created for the numerical solver. In this work, the optimization problem is solved using an interior points solver, however other solver types could also be used. After the solver has terminated, the feasibility of the solution is checked by verifying that the extremals of the lower level problem are indeed the maxima of the sensitivity functions on a predefined grid. Generally, the extremals of the lower level problem are not identified with exact accuracy and hence small absolute and relative tolerances `tol` are allowed, which are set to 0.01 throughout. Whether a solution is accepted or rejected in the presented algorithm then depends on the magnitude of the constraint violation γ and on the error in the identified peaks γ_i. Further checks based on constraint qualification could easily be implemented as well. If the desired conditions hold, the solution is returned. If either of the conditions are not satisfied, the initial guess is adapted. Within the scope of this work, this adaptation entails to resupply the optimizer with an initial guess informed by the location of the actual peak frequencies ω_l for a given p.

Algorithm 1: Bilevel optimization for PID tuning

Input : G, `tol`, Initial guess
Output : p, ω_l
Compute derivatives of $S(j\omega), T(j\omega), E_\mathrm{d}(j\omega)$
Construct problem and exact gradients
while $\gamma >$ *tol* **do**
  Solve bilevel optimization problem
  Compute γ_i for all peaks $i \in [S, T, E_\mathrm{d}]$
  if $\gamma_\mathrm{i} >$ *tol* **then**
    Update initial guess ω_l
  end
end

4.2 Implementation details

The problem is formulated in the Julia programming language using the `ModelingToolkit` package for automatic generation of symbolic derivatives. The problem is solved using the `Optimize` package, interfaced with the interior-point solver iPOPT. The lower level stationarity conditions ($\nabla_{\omega_\mathrm{l}}\mathcal{L} = 0$) are relaxed by introducing slack-variables ϵ, of which the squared values are added to the objective function multiplied by a weight w of 10^5 for the sensitivity function S, and 10^4 for all other transfer functions. This is akin to the penalty function method introduced in Sec. 2.4, leading to an augmented objective function in Eq. 14

$$J = \sum_{\mathrm{d}=1}^{\mathrm{D}} \left(\Phi_\mathrm{d} HIE_\mathrm{d} + w_\mathrm{d}\epsilon_\mathrm{d}^2\right) + w_\mathrm{S}\epsilon_\mathrm{S}^2 + w_\mathrm{T}\epsilon_\mathrm{T}^2. \quad (16)$$

In Eq. 16 the subscripts refer to the associated transfer functions.

5. RESULTS

The following section presents the results of two case studies of increasing complexity. In the first case study the developed method is applied to a first order plus time delay (FOPTD) process, and to an unstable FOPTD process in the second. While the considered plant transfer functions are seemingly simple, these case studies are selected as they are particularly challenging for the proposed tuning, because the sensitivity functions can exhibit multiple peaks with equal magnitude which makes locating the correct peak frequency non-trivial. Furthermore, the case studies illustrate the application of the developed method to two different scenarios, where a reasonable initial controller designer can be improved in terms of performance in case 1, or in terms of robustness in case 2.

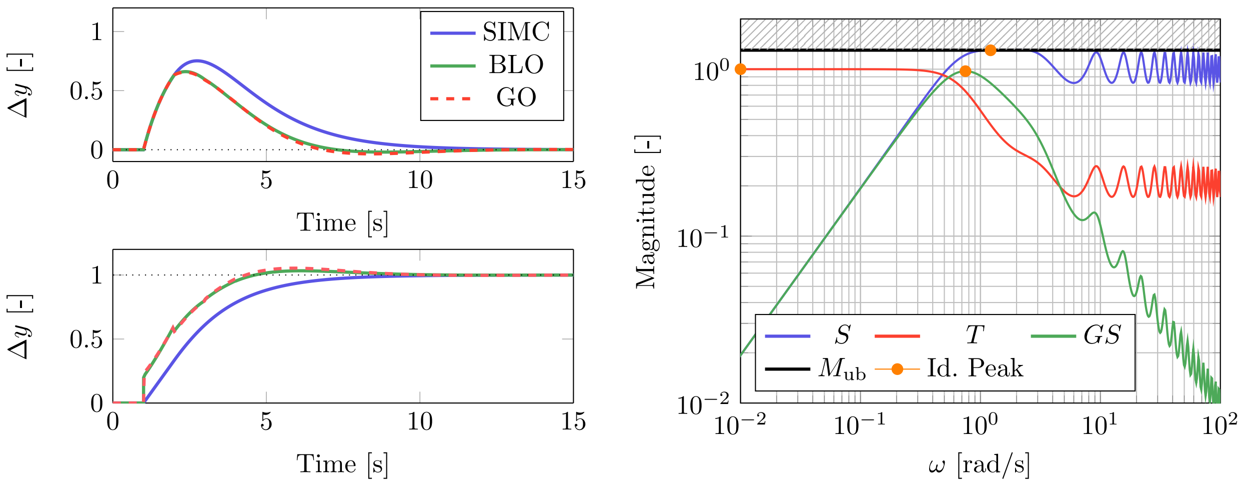
5.1 Case 1: Stable FOPTD

The first case study is concerned with the FOPTD system studied by Grimholt and Skogestad (2018b), and Turan et al. (2022).

$$G_1(s) = \frac{e^{-s}}{s+1} \qquad F(s) = \frac{1}{0.001s+1}, \quad (17)$$

where $F(s)$ is a filter to the controller, i.e. $K_\mathrm{f}(s) = K(s)F(s)$. For the sake of comparison with previous work, the same objective function is chosen as in aforementioned prior works, $J = \frac{1}{1.56}HIE_\mathrm{dy} + \frac{1}{1.42}HIE_\mathrm{du}$. Both the sensitivity and the complementary sensitivity functions are constrained in their magnitude such that $M_\mathrm{ST} \leq 1.3$, and only frequencies ω between 0.01 and 100 are considered. The supplied initial guess assumes all peak frequencies at 0.3 rad/s and $p = [0.33,\ 0.33,\ 0]$. This is the standard SIMC PI tuning rule in parallel form giving close to the desired robustness with $\tau_\mathrm{c} = 2\theta$. The derivative gain is constrained to be less or equal to 40 % of the integral gain. With these settings, the method identifies a suitable controller with one local solve in less than 0.1 seconds. The obtained tuning is shown in Table 1 together with those of Turan et al. and the SIMC rules. The herein presented bilevel optimization method is henceforth abbreviated as BLO and the global optimization approach employed by Turan et al. as GO. Table 1 further shows the IAE values for the considered tunings and the M_ST. The HIE values are not separately reported, as they can be inferred from the frequency domain performance, e.g. $HIE_\mathrm{dy} = \|S\|_\infty$.

The time domain performance of the controllers is shown in Fig. 2a, with Fig. 2b showing the sensitivity plot of the obtained controller together with the peaks identified by the optimization algorithm. It can be seen in Table 1 that the



(a) Disturbance rejection (upper), and setpoint tracking (lower)

(b) Frequency domain

Fig. 2. Comparison of time domain performances (a), and sensitivity function plot (b) for case 1.

Table 1. Tuning overview for case 1.

Method	Tuning	IAE_{dy}	IAE_{du}	M_{ST}
GO	[0.51,0.54,0.23]	2.2	2.07	1.3
BLO	[0.536,0.517,0.207]	2.14	2.06	1.306
SIMC	[0.33, 0.33, 0]	3	3	1.35

identified controller tuning is very similar to that obtained by Turan et al. using a global optimization approach, which itself are very similar to those Grimholt and Skogestad found using exact gradients on a fine frequency grid. As expected, the time domain performance for a unit step on the reference and on the output disturbance are almost identical for BLO and GO. Both show a faster settling time than the initial design in form of the standard SIMC tuning, but are also slightly less smooth in case of an output disturbance. Examining Fig. 2b, the optimizer manages to find the respective peaks of the sensitivity functions and to keep them bounded under M_{ub}. Upon closer inspection, the peak identified by the algorithm for $\|S\|_\infty$ lies slightly left of the actual peak. This gives rise to a constraint violation of 0.006, which could explain why the identified controller outperforms the one identified through a global optimization search in terms of IAE.

5.2 Case 2: Unstable FOPTD

In the second case study, the unstable FOPTD studied by Visioli (2001) is considered

$$G_2(s) = \frac{e^{-0.2s}}{s-1}. \tag{18}$$

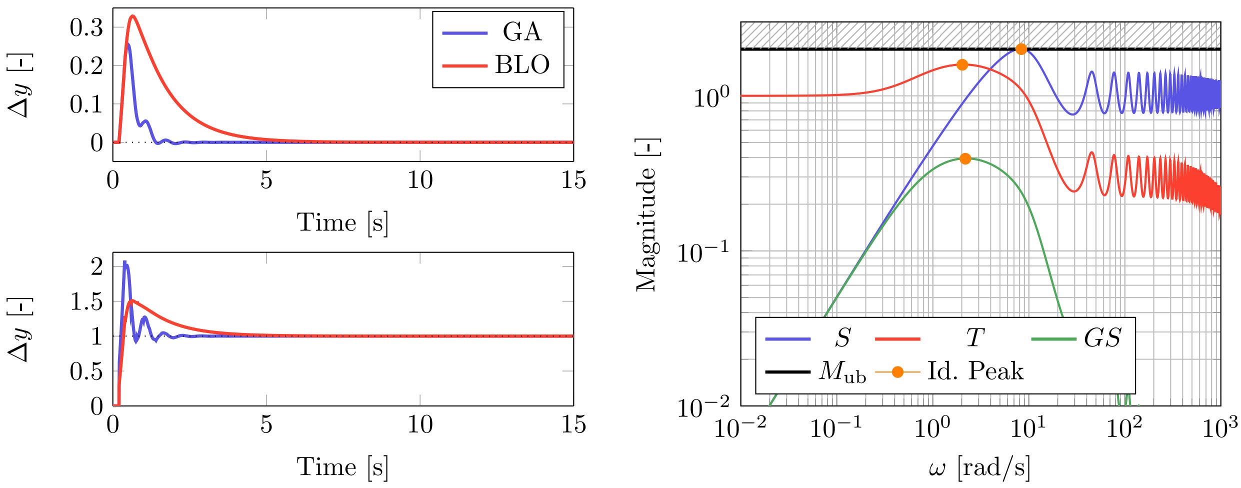
The author recommends two different PID-tunings, one for setpoint tracking and one for disturbance rejection. We herein compare the developed tuning with the setpoint tracking tuning, which has poor robustness with $M_{\mathrm{S}} = 4.1$. To improve on the robustness while retaining good setpoint tracking, the objective function $\eta = HIE_{\mathrm{dy}}$ is used while constraining M_{ST} to be below 2. The Bilevel optimization problem is solved by applying as initial guess Visiolis tuning and 10 rad/s for all frequencies. The derivative gain is constrained to be at most 15 % of the integral gain, and the same filter is applied as in the previous example. With this configuration, the algorithm finds a suitable controller with three local solves in typically less than 0.6 seconds combined. The solutions of the first two solves are rejected in accordance with the algorithm presented in Sec. 4 as the sensitivity function peak is wrongly identified. The identified controller tuning is shown in Table 2 together with that proposed by Visioli and the respective performance metrics. Note here that the controller parameters reported in the original publication of Visioli have been converted to the parallel form used in this work for consistency. Since the compared method is based on a genetic algorithm, it is further referred to as GA for brevity. Fig. 3 shows the time domain performance of both when subject to a unit step in the reference and the output, while b) shows the frequency dependent sensitivities of the identified PID controller. As expected, the less robust controller shows better performance, with the identified controller however also showing good results. As such, it shows a slightly larger settling time with a lower overshoot and less oscillatory response than GA in case of a setpoint change. Interestingly, the overshoot in case of an input disturbance is higher for the identified PID controller than for GA with again slightly larger settling times. Based on the performance metrics, increasing the robustness by a factor of two, i.e. from 4.16 to 2.01, comes at the expense of roughly doubling IAE_{dy}, with the increase in IAE_{du} being relatively higher but still in the same order of magnitude. Again, the obtained solution shows a minor constraint violation in M_{ST} within the required tolerance.

Table 2. Tuning overview for case 2.

Method	Tuning	IAE_{dy}	IAE_{du}	M_{ST}
BLO	[4.02,1.98,0.29]	1.02	0.12	2.01
GA	[6.23,8.54,0.46]	0.59	0.5	4.16

5.3 Discussion

The solutions obtained in the presented case studies are generally dependent on the initial guess, with most reasonable initial controller designs being improved upon. Other random guesses that are in the vicinity of a good design also achieve good performance, for example assuming uniform



(a) Disturbance rejection (upper), and setpoint tracking (lower)

(b) Frequency domain

Fig. 3. Comparison of time domain performances (a), and sensitivity function plot (b) for case 2.

frequencies and setting all values in p to either 5,6 or 7 respectively in case 2 leads to the same controller as the identified one. On the other hand, initial guesses that are relatively far from a desirable design will most likely converge to a less performant tuning. The same conclusion was reached on other conducted case studies, exhibiting larger delays or higher order functions. Furthermore, it was found that the performance of the found controllers greatly improved when imposing a constraint restricting the magnitude of K_d relative to K_i, as was done in the presented case studies. Depending on the initial guess, the optimizer would otherwise often find solutions where K_d exceeds K_i, which were found to be less performant when satisfying the same robustness requirements. Conclusively, the presented method is capable of improving initial designs both in terms of robustness and performance with little computation effort and fast solving times. While efficient, the underlying local optimization requires sensible initial guesses and is hence believed to be best suited to improve on existing controller designs obtained for example by established tuning rules. Applying expert knowledge to restrict the search space further improves the performance.

6. CONCLUSION

This work presents a novel approach for tuning PID controllers with robustness constraints based on bilevel optimization principles. Here, the optimality conditions of the sensitivity function peaks are embedded as constraints in the optimization problem, which entails minimizing an error function in the frequency domain. A key element of the method is the automatic symbolic differentiation of the objective and constraint functions which facilitates automated representation of the optimization problem and allows for efficient numerical solving due to the availability of exact gradients. The main benefit compared to previous work is then the lower computational complexity and higher adaptability of the approach. This however comes at the expense of guaranteed constrain satisfaction, for example when the sensitivity functions exhibit multiple peaks. The method is applied to two case studies which exhibit this challenging behavior, namely a stable and a unstable FOPTD system. In both cases, constraint satisfaction is achieved within low tolerances, and well performing controllers are found in the order of roughly 0.1 seconds.

REFERENCES

Åström, K. and Hägglund, T. (2006). *Advanced PID control*. ISA - The Instrumentation, Systems and Automation Society.

Balchen, J. (1958). *A performance index for feedback control systems based on the fourier transform of the control deviation*. Acta polytechnica Scandinavica. Norges tekniske vitenskapsakademi.

Garpinger, O., Hägglund, T., and Åström, K.J. (2014). Performance and robustness trade-offs in PID control. *Journal of Process Control*, 24(5), 568–577.

Grimholt, C. and Skogestad, S. (2018a). Optimal PI and PID control of first-order plus delay processes and evaluation of the original and improved simc rules. *Journal of Process Control*, 70, 36–46.

Grimholt, C. and Skogestad, S. (2018b). Optimization of fixed-order controllers using exact gradients. *Journal of Process Control*, 71, 130–138.

Sinha, A., Malo, P., and Deb, K. (2018). A review on bilevel optimization: From classical to evolutionary approaches and applications. *IEEE Transactions on Evolutionary Computation*, 22(2), 276–295.

Skogestad, S. (2003). Simple analytic rules for model reduction and PID controller tuning. *Journal of Process Control*, 13(4), 291–309.

Turan, E.M., Kannan, R., and Jäschke, J. (2022). Design of PID controllers using semi-infinite programming. In Y. Yamashita and M. Kano (eds.), *14th International Symposium on Process Systems Engineering*, volume 49 of *Computer Aided Chemical Engineering*, 439–444. Elsevier.

Visioli, A. (2001). Optimal tuning of PID controllers for integral and unstable processes. *IEE Proceedings - Control Theory and Applications*, 148(2), 180–184.

Ziegler, J.G. and Nichols, N.B. (1942). Optimum settings for automatic controllers. *Journal of Fluids Engineering*, 64(8), 759–765.