Genetic PI Controller Tuning to Emulate a Pole Assignment Design

Marco Paz Ramos ∗ Axel Busboom ∗ Andriy Slobodyan ∗∗

∗ Department of Engineering and Management, Munich University of Applied Sciences, Munich, Germany
∗∗ Business Unit Catalysts, Clariant AG, Munich, Germany

Abstract: Tuning of proportional-integral (PI) and proportional-integral-derivative (PID) controllers continues to be a current topic, as the control needs in industry are broad and diverse. Although PID controllers have been the predominant controller type for several decades, there are still opportunities to improve the performance of both the installed base and newly deployed controllers. One of the main challenges is a reliable and easy tuning that can be performed by an operator on site. From a computer science point of view, the problem of tuning PID controllers qualifies as a nondeterministic polynomial-time hard problem (NP-hard). Genetic algorithms are a heuristic approach to approximate this type of problem, and the continued growth of computational capacity makes them increasingly more viable. In this work, we present a PID tuning architecture using a genetic algorithm, which incorporates in its fitness function an emulation of pole assignment and implicit cancellation of the additive dynamics of zeros in a closed loop, in addition to reducing discretization losses. The tuning is performed with a paradigm different from the one usually applied in the literature. Rather than simply minimizing the error between the process variable and the setpoint, the error between the process variable and an ideal response curve associated with a desired pole assignment is minimized. This approach provides better control over closed-loop performance.

Keywords: PID, PI, Genetic Algorithms, Pole Assignment.

1. INTRODUCTION

The field of control engineering is very broad and diverse, and new contributions are continuously made to this field. Despite these advances, the de facto standard in the process control industry for single-loop controllers, e.g., for temperature, flow, or level control, is still proportional-integral-derivative (PID) controllers. It is estimated that approximately 90% of the control loops in industry are PID type (Åström and Hägglund, 2001; Knospe, 2006).

Since its inception, probably the most significant challenge in implementing PID controls has been that of parameter tuning. Early operators did not have the theoretical and practical training necessary for accurate tuning, which led to the introduction of simple heuristics such as those proposed by Ziegler and Nichols and many others (O'Dwyer, 2009). Although understanding of the dynamics of PID loops has increased, the practical implementation of PID control often continues to be suboptimal. It is estimated that up to one third of the installed loops are manually tuned and that at least one in four uses the manufacturer’s default gain settings (Ender, 1993). Analytical tuning requires mathematical modeling of or data from the plant in either the time or frequency domain. At the higher end of the spectrum is exhaustive or mechanistic mathematical modeling, which requires detailed and significant consideration of the physical and/or chemical laws involved. This approach usually results in complex models that may require custom controllers. Halfway to the simplest heuristic tuning, there is the alternative of designing controllers using black box models, i.e., functional relationships between the inputs and outputs of a system that represent the main characteristics of the system without being based on physical relationships (Zhang, 2010).

Many plants to be controlled by PID controllers can reasonably be approximated as first- or second-order linear models. If such approximate models are available, the calculation of gains by pole assignment is feasible. Beyond the representativeness of the chosen model, there are several reasons why the performance of the practical implementation may differ from that of the theoretical control design. Aspects that may cause discrepancies between design and implementation performance include:

(a) the additive dynamics of zeros when closing the loop;
(b) effects of discretization when using digital controllers (unless this is considered in the design from the beginning);
(c) differences in the algorithm implementation between the design and the actual industrial controller;
(d) discrepancies between digital algorithms and configuration errors when selecting time bases.

Since the tuning task only involves choosing three parameters, it may seem simple. However, like all nonconvex optimization problems, tuning the PID parameters for an...
accurate and stable closed-loop control becomes an NP-hard problem (Somefun et al., 2021). The complexity of the tuning task can be approached as a multi-objective optimization complying with a Pareto front (Guenounou et al., 2012) and 4) it allows circumscirbing validity zones for standard structures, as is the case for tuning tables (Visioli, 2001).

In the literature on genetic tuning of PID controllers, error criteria such as Integral of the Squared Error (ISE), Integral of the Absolute magnitude of the Error (IAE), or Integral of Time multiplied by Absolute Error (ITAE) (Borase et al., 2021) are frequently used. These have the advantage of directly using the tracking error with respect to the setpoint given during the control. However, a disadvantage is that the desired closed-loop performance cannot be precisely established beyond the error minimization itself. For example, with ITAE, the genetic algorithm will aim to make the control response as fast as possible. In practice, however, other considerations such as actuation saturation must be taken into consideration. Tuning PI controllers by pole assignment can be useful for establishing specific closed-loop responses, but even the additive dynamics of zeros can slightly change the closed-loop response from an originally desired one.

This article presents a proposal to tune PI and PID controllers with genetic algorithms. It applies an optimization criterion to obtain a specific closed-loop response as in tuning by pole assignment, but implicitly mitigates the effects of the additive dynamics of zeros and of discretization.

2. POLE ASSIGNMENT DESIGN

PI tuning by pole assignment is well understood (Aström and Hägglund, 1995). This section deals with the approach to transfer functions discussed in (Paz et al., 2017). Assume we have a first-order plant of the form

\[ G_1(s) = \frac{K}{\tau s + 1} \]  

where \( K \) is the open loop gain and \( \tau \) is the time constant. Now we close the loop with a PI controller (Aström and Hägglund, 1995) of the form

\[ G_c(s) = \frac{K_c T_i s + K_c}{T_i s} \]  

where \( K_c \) is the proportional gain of the controller and \( T_i \) is the integral time. With negative feedback, the following closed-loop transfer function is obtained

\[ G_{CL1}(s) = \frac{K K_c (T_i s + 1)}{s^2 + \frac{1}{\tau}(1 + K K_c) s + \frac{K K_c}{\tau^2}}, \]  

which is similar to the canonical form of a second-order transfer function, but differs by the presence of a zero in the numerator

\[ G_{SO}(s) = \frac{\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2}. \]  

Here, \( \xi \) is the damping coefficient and \( \omega_n \) is the undamped natural frequency. From the denominator, we can calculate the approximate maximum overshoot \( M_p \) when \( \xi \) is in the range between 0 and 1

\[ M_p \approx \frac{1}{\sqrt{4\pi^2 + 1}}, \]  

and the settling time \( t_s \) defined within an error band of 1%, (Franklin et al., 1998)

\[ t_s \approx \frac{4.6}{\xi\omega_n}. \]  

If \( K_c \) and \( T_i \) are known, we can calculate the settling time \( t_s \) and the maximum overshoot \( M_p \) for the closed loop from (5) and (6).

To tune the controller gains, we specify a desired closed-loop settling time \( t_{sd} \) and maximum overshoot \( M_{pd} \), and solve (5) and (6) for the desired damping coefficient \( \xi_d \) and undamped natural frequency \( \omega_{nd} \):

\[ \xi_d \approx \frac{\ln^2(M_{pd})}{\pi^2 + \ln^2(M_{pd})}, \quad \omega_{nd} \approx \frac{4.6}{\xi_d t_{sd}}. \]  

Finally, by comparing the denominators of equations (3) and (4) we get

\[ K_c = \frac{2\tau \xi_d \omega_{nd} - 1}{K}, \quad T_i = \frac{K K_c}{\tau (\omega_{nd})^2} \]  

for the parameters of the PI controller.

If the process can be approximated as an overdamped second-order system (\( \xi > 1 \))

\[ G_2(s) = \frac{K}{(\tau_1 s + 1)(\tau_2 s + 1)}, \]  

when the PI control of equation (2) is applied to the plant of (9), the closed-loop transfer function becomes

\[ G_{CL2}(s) = \frac{K K_c (\tau_1 s + 1)}{s^3 + \left(\frac{\tau_1 + \tau_2}{\tau_1 \tau_2}\right)s + \left(\frac{K K_c + 1}{\tau_1 \tau_2}\right)s + \frac{K K_c}{\tau_1 \tau_2}}. \]  

Following the same scheme of equalizing denominators of the controlled plant and the desired plant, it is necessary to add a third pole to the desired denominator of the form

\[ D(s) = (s + p_1)(s^2 + 2c\xi\omega_{nd}s + \omega_{nd}^2). \]  

As in the design for the first-order plant, the desired second-order closed-loop component can be constructed by choosing the desired settling time \( t_{sd} \) and the desired maximum closed-loop overshoot \( M_{pd} \). Term-by-term matching of (11) and the denominator of (10) yields

\[ p_1 = \frac{\tau_1 + \tau_2}{\tau_1 \tau_2} - 2\xi_c\omega_{nd}. \]  

However, it is not sufficient to arbitrarily choose the value for the desired settling time \( t_{sd} \) to construct the second-order component of the desired polynomial as this may
lead to a polynomial $D(s)$ that is not asymptotically stable. To avoid this, the following necessary and sufficient condition for the settling time must be fulfilled (Paz et al., 2017):

$$t_{sd} > \frac{9.2(\tau_1 \tau_2)}{\tau_1 + \tau_2}.$$  \hspace{1cm} (13)

Following the same equalization process as above, the controller gains for this case can be obtained as

$$K_c = \frac{\omega_n^2\tau_1\tau_2 + 2p_1\xi\omega_n\tau_1\tau_2 - 1}{K},$$  \hspace{1cm} (14)

$$T_i = \frac{K_c K}{\tau_1\tau_2\omega_n^2p_1}.$$  \hspace{1cm} (15)

3. GENETIC ALGORITHMS

Evolutionary computation is a family of computational optimization techniques inspired by natural evolution (Eiben and Smith, 2015). Examples include genetic algorithms, evolutionary programming, genetic programming, memetic algorithms, or ant systems.

Genetic algorithms are probably the most representative technique of evolutionary computation and have been adopted in the field of automatic control for a long time (Wang et al., 2003). Genetic algorithms are inspired by the exchange of genetic information in evolution in nature. Although optimization techniques based on gradient descent are very popular in machine learning due to their efficiency, in some optimization problems where a balance between exploration and exploitation is required, evolutionary optimizers can be more effective and have multiple advantages: Genetic algorithms are an inherently parallel method (Goldberg, 1988), which can maintain a population of possible solutions at a given time and whose final population is part of the Pareto front (Wei and Sööker, 2015). The use of genetic algorithms is viable when performing exploration can contribute to the search for a better result. Furthermore, genetic algorithms can support cost functions with complex rules and restrictions since they are related to an execution approach rather than a specific calculation. Finally, genetic algorithms are good when the objective function has high modality, i.e., many local optima (Mirjalili, 2019).

The optimization of PID gains in a discrete space can be understood as an NP-hard problem. Depending on the requirements, there may be several acceptable solutions for tuning a controller; i.e., the tuning problem can then be interpreted as a combinatorial one. This makes it appropriate to use a genetic algorithm, which beyond the search for a valley or crest, searches for various combinations that may be viable for the joint satisfaction of a set of restrictions.

4. GENETIC PID WITH EMULATION OF POLE ASSIGNMENT

Genetic algorithms can be applied to the tuning of PID controllers in different ways. In this work, a parametric optimization architecture is proposed in which a genetic algorithm incorporates a black-box model that abstracts the dynamics of a given process. Controller tuning is approached as a combinatorial optimization problem with the objective of choosing the controller parameters in such a way that the desired response of the internal model is achieved. Fig. 1 shows the architecture of the proposed emulative pole assignment scheme for PID controllers.

Fig. 1. Architecture of the proposed emulative pole assignment scheme.

The individual steps in the diagram are explained below.

1) A table is generated with random candidate values of the controller gains. In the terminology of evolutionary computation, this first set of potential solutions is called a population and each ordered triad of parameters

$$C_{k,n} = [K_{c,n}, T_{i,n}, T_{d,n}]$$  \hspace{1cm} (16)

is called an individual or chromosome where $k$ is the generation and $n$ is the number of individuals in a given population.

2) The parameters of this initial population are applied to a dependent PID in velocity form (Åström and Hågglund, 1995), which is common in industrial digital controllers

$$u(t) = u(t-1) + K_c [e(t) - e(t-1)] + \frac{T_i}{T} e(t)$$

$$+ \frac{T_d}{T} (e(t) - 2e(t-1) + e(t-2))$$

where $u$ is the control signal, $e$ is the error, $T$ is the sampling period and $T_d$ is the derivative time. Each element of the population is evaluated sequentially.

3) Using equation (7) from a desired overshoot $M_{pd}$ and desired settling time $t_{sd}$, a transfer function of the
desired dynamics is calculated. This transfer function is discretized using a sampling period $T$, then transformed from a differential equation to a difference equation that synthesizes the expected response. From this, the expected step response is vectorized within a time window that contains at least the desired settling period $t_{sd}$.

4) Each chromosome of the population of generation $k$ is evaluated within a time window that contains the desired settling time $t_{sd}$ and the response of the controlled plant is vectorized. The black box model, depending on the case, corresponds to a discretized model with a zero-order hold (ZOH) of equations (1) or (12), which is implemented as a recursive difference equation (Franklin et al., 1998).

5) In genetic algorithms, the objective function is referred to as the “fitness function”. Here it is the absolute error of the standard response of the desired pole assignment and the response obtained with the $n^{th}$ test chromosome

$$
ff_n = \sum |yd_n - yg_n|
$$

where $ff_n$ is the fitness function of an individual $n$, $yd_n$ is the vector of the desired output of the controlled plant, and $y_n$ is the output using the $n^{th}$ test chromosome. We have chosen the absolute error as opposed to a squared error, as the squared error tends to focus too much on small time intervals with the largest error, while the absolute error tends to lead to an overall closer approximation of the desired closed-loop response.

6) Once all chromosomes in a population $k$ have been labeled with their corresponding fitness functions, we proceed to a “tournament” (Goldberg and Sastry, 2007; Kramer, 2017), the size of which is determined as a percentage of the population size (10% of the population size is used for this design to promote genetic diversity). Through the tournament, pairs of potential parents are selected to carry out the reproductive process.

7) With the selected parents, a bit-level crossover is performed (Goldberg and Sastry, 2007; Kramer, 2017). A random cutoff point is chosen along the chromosomes, where each of the parents’ chromosomes is separated into two parts. The resulting sections are now recombined in such a way that the lower part of one parent joins the upper part of the other to form a place for a child, while the complementary parts form the other. In this case, each pair of parents begets a pair of children, which are partially different from their parents and will become part of the $k + 1^{st}$ population.

8) Once the same number of offspring as parents have been assembled in a new population (to maintain ecological stability), mutations are made in a small percentage of individuals (Goldberg and Sastry, 2007; Kramer, 2017), i.e., a randomly chosen bit of a given chromosome is inverted. Mutation is an important operator to maintain genetic diversity on a generational scale.

9) Finally, the population obtained in the $k+1^{st}$ generation will entirely replace the previous generation.

10) The Holland’s Schema Theorem (Holland, 1992) states that segments of the chromosome more fit than average will generate copies on an exponential scale in subsequent generations. Therefore, the fittest individual of each generation will provide a closed-loop performance $y_k$ increasingly similar to the desired $y_d$ and the error will decrease over the generations as far as possible, respecting the closed-loop structural capacity.

5. PI CONTROLLER GENETIC TUNING TESTS

This section presents some examples to compare the performance of the pole assignment method described in Section 2 with the genetic tuning algorithm. We assess how similar the closed-loop temporal response for both methods is to an ideal closed-loop response that meets a desired settling time $t_{sd}$ and a maximum overshoot $M_{opd}$ according to (4) and (11). The difference with respect to the reference curve is quantified by the sum of the absolute errors, described as the fitness function $ff_n$ of equation (18). For all test cases, the genetic algorithm used was configured to compute 50 generations with a population size of 500, a chromosome size equal to 2 ($K_c$, $T_i$), and a mutation index of 6%. The resolution of each gene is 53 bits as before the decoding process, integers are used in significant IEEE double precision (IEEE 754-2019). These hyperparameters were chosen heuristically as a compromise between convergence and applicability for the set of cases presented in this section.

Fig. 2 shows an example of the step responses for a first-order plant of the form (1) with an open-loop gain of 1.2 and a time constant of 6.9. For this case, $t_{sd} = 3\tau$ and a $M_{opd} = 0.1$ have been used, so from (8) we get $K_c = 1.722$ and $T_i = 2.1195$. The tuning of the genetic algorithm yields gains $K_c = 0.9766$ and $T_i = 2.8253$. The sampling period is $T = 1$ for both tuning methods. It can be seen that the tracking of the target pole assignment curve is closer with the genetic tuning method.

![Fig. 2. Discrete PI performance tuned by a genetic algorithm to emulate a non-zero dynamic pole assignment for a first-order plant (ZOH, T=1).](image-url)
genetic tuning than with pole assignment, which shows a better capability of the proposed method to meet the design requirement. On average, the cumulative absolute error of the genetic algorithm was 74.4% lower than with pole assignment for these first-order cases. The * symbol in the table indicates that for the given combination of $M_p$ and $t_{sd}$, the closed loop of the pole assignment method was unstable. While this problem could be easily fixed by decreasing the sampling period $T$, the experiment is meant to demonstrate the limits of both tuning methods. It can be seen that the genetic algorithm was able to find a viable solution even in the situations where the conventional pole assignment yielded an unsolvable solution.

Table 1. Sum of the absolute difference between curves, $f$, first-order plant (1), $T = 1$, time window length 35.

<table>
<thead>
<tr>
<th>$M_p$</th>
<th>p.a.</th>
<th>g.t.</th>
<th>$M_p$</th>
<th>p.a.</th>
<th>g.t.</th>
<th>$M_p$</th>
<th>p.a.</th>
<th>g.t.</th>
<th>$M_p$</th>
<th>p.a.</th>
<th>g.t.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>1.15</td>
<td>0.998</td>
<td>0.071</td>
<td>0.098</td>
<td>0.332</td>
<td>0.173</td>
<td>0.093</td>
<td>0.146</td>
<td>0.292</td>
<td>0.903</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>2.126</td>
<td>2.283</td>
<td>2.012</td>
<td>1.741</td>
<td>1.853</td>
<td>1.804</td>
<td>1.785</td>
<td>1.857</td>
<td>1.841</td>
<td>1.631</td>
<td></td>
</tr>
<tr>
<td>0.4</td>
<td>2.885</td>
<td>2.898</td>
<td>2.871</td>
<td>2.840</td>
<td>2.840</td>
<td>2.840</td>
<td>2.840</td>
<td>2.840</td>
<td>2.840</td>
<td>2.840</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>2.635</td>
<td>2.635</td>
<td>2.635</td>
<td>2.635</td>
<td>2.635</td>
<td>2.635</td>
<td>2.635</td>
<td>2.635</td>
<td>2.635</td>
<td>2.635</td>
<td></td>
</tr>
</tbody>
</table>

The similarity between the pole assignment tuning and the desired target curve could be increased by adding a feedforward block for closed-loop zero cancellation, but this is difficult to implement on commercial off-the-shelf controllers. The proposed genetic tuning approach emulating pole assignment with zero cancellation does not require modifications to the standard PID structure.

Additional factors may not be considered in the practical deployment of the controller. As mentioned above, deviations due to discretization must be considered during the design stage, particularly when the dynamics of the process are in the same order of magnitude as the sampling period. Other aspects that may influence the practical performance of the tuning include (a) sensor delays, (b) noise, (c) loss of accuracy, (d) actuator saturation, and (e) actuator dynamics, e.g., valve motion. Each of these aspects can change the performance of the control loop from the desired behavior. Three of them can be incorporated directly into the genetic controller tuning scheme: 1) noise as an observable percentage of the steady-state response of the plant, 2) actuator saturation (e.g., valve limits), and 3) actuator dynamics.

As with the first-order example, two parameters were varied. In this case, the variation on the desired maximum overshoot $M_{pd}$ was maintained and $\tau_2$ was chosen as the parameter to be varied. The values of $\tau_1$, $K$ and $t_{sd}$ were maintained from the example above. As in the previous case, it can be seen from Table 2 that the cumulative error is always lower for genetic tuning than with pole assignment (58.4% in average across all cases).

Table 2. Sum of the absolute difference between curves, $f$, second-order plant (9), $T = 5$, time window length 120.

<table>
<thead>
<tr>
<th>$M_p$</th>
<th>p.a.</th>
<th>g.t.</th>
<th>$M_p$</th>
<th>p.a.</th>
<th>g.t.</th>
<th>$M_p$</th>
<th>p.a.</th>
<th>g.t.</th>
<th>$M_p$</th>
<th>p.a.</th>
<th>g.t.</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>2.220</td>
<td>2.154</td>
<td>0.699</td>
<td>1.566</td>
<td>1.566</td>
<td>1.080</td>
<td>0.422</td>
<td>0.618</td>
<td>0.339</td>
<td>0.473</td>
<td></td>
</tr>
<tr>
<td>0.1</td>
<td>2.600</td>
<td>1.748</td>
<td>0.941</td>
<td>1.471</td>
<td>0.821</td>
<td>1.332</td>
<td>0.757</td>
<td>1.265</td>
<td>0.593</td>
<td>1.283</td>
<td></td>
</tr>
<tr>
<td>0.25</td>
<td>2.840</td>
<td>1.144</td>
<td>1.009</td>
<td>1.086</td>
<td>0.919</td>
<td>1.086</td>
<td>0.919</td>
<td>1.086</td>
<td>0.919</td>
<td>1.086</td>
<td></td>
</tr>
</tbody>
</table>

1 A heuristic way to calculate the approximate settling time in overdamped systems is to multiply the sum of both time constants by 3.5.

2 As a rule of thumb, a value of one second in the time constant of the response of the valve per inch of the valve size is often used in practice.
For both controllers, additive noise is simulated with zero mean and standard deviation $1 \times 10^{-3}$, actuator limits of 0 and 1 are considered, and an actuator time constant of 3 is simulated. These effects are taken into account in the training of the genetic algorithm. While the use of anti-windup, which temporarily turns off the accumulation on the integral gain while the actuator is at a limit, is an industry standard, the inclusion of saturation in the genetic optimization aims at taking constraints into account already in the controller design and designing an intrinsically less aggressive controller if necessary. Fig. 4 shows a similar initial slope of the step response due to actuator saturation for both approaches. However, genetic tuning provides weaker oscillations and an overall error 20% lower compared to that of pole assignment, underscoring the point that the algorithm can be readily used to account for non-ideal effects in the plant behavior.

6. CONCLUSION
This study presented a proposal to tune PI controllers with genetic algorithms. Instead of a cost function based on the tracking error at the setpoint, we use one based on the minimization of the error between a desired and feasible step response (designed to achieve both a certain settling time and a maximum overshoot) and the actual closed-loop step response. This emulated pole assignment approach yields a closed-loop response that is closer to the target curve than conventional pole assignment. The method is also capable of accounting for practical aspects, such as noise or actuator saturation and dynamics. On the other hand, the method proposed in this work could encounter limitations if the quality of the approximate black box model is not adequate. Finally, the proposed structure can be extended to plants of different types and even incorporate an operator interface that facilitates the intuitive choice of performance criteria.

REFERENCES