Computers and Chemical Engineering 173 (2023) 108211

Contents lists available at ScienceDirect

Computers and Chemical Engineering

journal homepage: www.elsevier.com/locate/cace

Bayesian Optimization for automatic tuning of digital multi-loop PID controllers

João P.L. Coutinho, Lino O. Santos, Marco S. Reis *

Univ Coimbra, CIEPQPF, Department of Chemical Engineering, Rua Sílvio Lima, Pólo II - Pinhal de Marrocos, 3030-790 Coimbra, Portugal

ARTICLE INFO

Keywords:
Multi-loop PID tuning
Automatic tuning
Bayesian optimization
Internal model control
Sequential loop closing

ABSTRACT

In recent years, the use of Bayesian optimization (BO) for efficient automatic tuning of general controller structures through iterative closed-loop experiments, has been attracting increasing interest. However, its potential for tuning interactive multi-loop PID controllers in Multi Input Multi Output (MIMO) processes remains largely unexplored. Even though the optimization domain greatly affects closed-loop performance and safety, it is usually defined manually, through expert knowledge or experimentation. This paper presents a novel systematic methodology for defining the optimization domain for automatic multi-loop PID tuning using BO. Sequential loop closing, system identification and tuning relations are used to constrain the bounds on controller parameters to meaningful ranges, including gains and sampling times. This provides an effective way to improve the convergence of BO and secure process safety during closed-loop experiments, without requiring a MIMO process model or extensive prior knowledge. The methodology can be applied "as is" to single-loop PID controllers.

1. Introduction

Multi-loop, or decentralized, Proportional Integral Derivative (PID) controllers remain the most widely used solution for regulatory control of Multiple Input Multiple Output (MIMO) industrial processes. This is due to their flexibility, simplicity, and inherent failure tolerance, when compared to centralized control structures. Moreover, their flexibility and performance can be further improved with commercially available features such as anti-windup, setpoint weighting and feed-forward compensation (Åström and Hägglund, 2006). Even when more advanced supervisory controllers are adopted, e.g. Model Predictive Control (MPC), PID controllers are still usually present in the lower regulatory layers.

A vast number of single-loop PID controller tuning methods are available, including model-based analytic rules (O'Dwyer, 2009), ultimate frequency methods (Ziegler and Nichols, 1942; Åström and Hägglund, 1984) and optimization-based tuning (Åström and Hägglund, 2006). When interactions in MIMO processes are significant, tuning each decentralized controller independently with any of these methods can lead to poor performance or, worse, closed-loop instability of the multivariable system, even when each individual loop is stable (Ogunnaike and Ray, 1994). Several methods have been proposed for tuning decentralized PID controllers considering loop interactions, generally categorized into model-based approaches, such as detuning (Luyben, 1986; Chien et al., 1999; Chen and Seborg, 2002), sequential loop closing (Hovd and Skogestad, 1994) and independent design (Hovd and Skogestad, 1993; Vu and Lee, 2010) or model-free sequential (Loh et al., 1993; Shen and Yu, 1994) and decentralized (Halevi et al., 1997) relay autotuning. Despite the solutions reported in the scientific literature, optimally tuning decentralized PID controllers remains a difficult and complex task. The industrial practice often consists of sequential loop closing followed by heuristic and conservative detuning, which is a complex and time consuming process (Dittmar et al., 2012).

In recent years, the focus has shifted to treating multi-loop PID design as a, possibly constrained, optimization problem, solved off-line using a MIMO process model (Sumana and Venkateswarlu, 2010; Xue et al., 2010; Dittmar et al., 2012; Euzebio et al., 2021). While robustness can be taken into account by considering appropriate constraints (Dittmar et al., 2012; Euzebio et al., 2021), such is done to ensure worst-case closed-loop stability, not performance, in the face of process variations or modeling errors. Ultimately, closed-loop performance is dependent on an accurate MIMO process model, which can be difficult or expensive to identify for non-linear multivariable processes. Alternatively, data-driven optimization-based tuning methods can be used, that do not explicitly require a process model. Instead they optimize controller parameters to either match a reference closed-loop response based on collected open-loop data (Campi et al., 2002;

* Corresponding author.
E-mail address: marco@eq.uc.pt (M.S. Reis).

https://doi.org/10.1016/j.compchemeng.2023.108211
Received 1 November 2022; Received in revised form 17 February 2023; Accepted 3 March 2023
Available online 6 March 2023

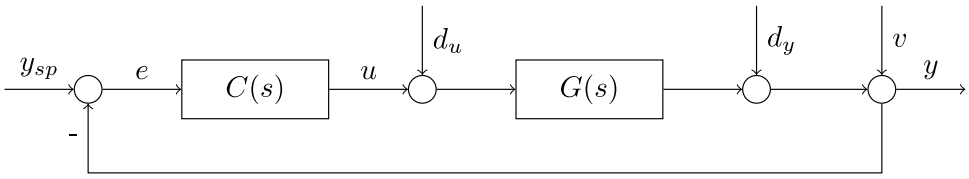

Fig. 1. MIMO feedback control system.

Rojas et al., 2012; Campestrini et al., 2016), or to minimize a desired objective based on gradient information obtained through iterative closed-loop experiments (Hjalmarsson, 2002). However, the former approach can be hard to extend to arbitrary performance objectives and constraints while the latter consists of a local search method which may be slow to converge or require extensive experimentation.

Recently, Bayesian optimization (BO) (Brochu et al., 2010; Shahriari et al., 2016; Greenhill et al., 2020) has stood out as a sample-efficient data-driven global optimization method for automatic controller tuning, including PID (Schillinger et al., 2017; Khosravi et al., 2019; Fiducioso et al., 2019; Konig et al., 2021; Khosravi et al., 2022a,b), MPC (Piga et al., 2019; Lu et al., 2021; Sorourifar et al., 2021; Makrygiorgos et al., 2022) and other controllers (Marco et al., 2017; Fröhlich et al., 2019; Neumann-Brosig et al., 2020). The performance of BO critically depends on a proper definition of the optimization domain. A poor choice can lead to instability or cause slow convergence, steeply increasing the experimental tuning cost. Hitherto, this domain is defined based on the information obtained from previous experiments on a simulated or physical process, expert knowledge, or using grid search, which is infeasible on a real system. The requirement of manual tuning or previous trial-and-error experiments presents a disadvantage to the use of BO for automatic multi-loop PID controller tuning, especially for the case of multivariable processes, where manually defining a design space could be difficult to do because of the existence of interactions among variables. Therefore, in this work, we present a systematic methodology that combines sequential loop closing, system identification and model-based tuning relations, to define a relevant optimization domain for multi-loop PID controller tuning with BO. This methodology improves the tuning efficiency by constraining the optimization search in a relevant domain and reduces the need for trial-and-error experiments or expert knowledge to apply BO for controller tuning. Due to its sequential nature, the proposed methodology can also be applied "as is" for single-loop PID controller tuning.

The remainder of the article is organized as follows. Section 2 presents the problem definition in detail. Section 3 introduces the proposed methodology for decentralized PID controller tuning using Bayesian Optimization, which is applied to a simulation case study in Section 4. Finally, Section 5 summarizes the conclusions of this work, and points some future research directions.

2. Problem formulation

2.1. MIMO system

Consider the MIMO feedback control system with n multi-loop controllers, represented in Fig. 1.

C is a diagonal matrix, $C = \text{diag}\{C_1(s), \dots, C_n(s)\}$, and $G(s)$ is a $n \times n$ process transfer function matrix:

$$G(s) = \begin{bmatrix} G_{11}(s) & \cdots & G_{1n}(s) \\ \vdots & \ddots & \vdots \\ G_{n1}(s) & \cdots & G_{nn}(s) \end{bmatrix}. \tag{1}$$

y_{sp}, y and u denote the n-dimensional vectors of setpoints, measured process outputs and process inputs, respectively. The system is subject to input, d_u, and output, d_y, disturbances as well as measurement noise, v. In this work, we consider each decentralized controller C_i as a discrete PI controller defined using the position form and setpoint weighting:

$$u(k) = u_0 + K_P(by_{sp}(k) - y(k)) + K_I \Delta t \sum_{j=1}^{k}(y_{sp}(j) - y(j)), \tag{2}$$

where k is the sampling instant, Δt the controller sampling time, u_0 the bias signal, K_P, K_I and b are the proportional gain, the integral gain and proportional setpoint weight, respectively. To account for actuator constraints, the saturated controller output, $sat(u(k))$, is limited between minimum, u_{min}, and maximum, u_{max}, values. Because we are considering a PI controller with the position form, conditional integration, or clamping, is added as an anti-windup measure (Åström and Hägglund, 2006).

In addition to the selection of the controller gains, the parameters b and Δt need to be specified. Although b is usually set to 1, previous works have shown that the use of setpoint weighting, with $b = 0$, can reduce loop interactions and improve decentralized controller performance (Loh et al., 1993; Chien et al., 1999; Chen and Seborg, 2002). This happens to be so because in a decentralized control structure a setpoint change in one loop acts as a disturbance to other loops. Therefore, in this work we will consider b as an additional tuning parameter, constrained to fall between 0 and 1.

The selection of a proper value for Δt in digital PID controllers remains an open-ended problem, usually solved with manual tuning using heuristics. However, this parameter affects both closed-loop performance and stability. A very large value compared to the process dominant time constant can lead to instability (Ogunnaike and Ray, 1994). On the other hand, a very small value can lead to excessive controller action due to measurement noise and increased actuator wear and tear. In this work, we propose to optimize Δt simultaneously with the remaining PI controller parameters. This is done by applying BO on physical closed-loop experiments which does not require the same Δt between iterations or for different loops. We assume, however, that controlled variables are monitored with a faster, but fixed, sampling rate during these experiments, and that Δt can be treated as a continuous decision variable with negligible rounding error.

2.2. Performance evaluation

In this work, we quantify the performance of each loop with the Integrated Absolute Error (IAE), to assess the tracking performance, and the Total Variation of the controller output (TV), to measure the control effort, using discrete approximations for both:

$$IAE = \int_0^\infty |y_{sp}(t) - y(t)|dt = \sum_{k=0}^{K}|y_{sp}(k) - y(k)|\Delta t \tag{3}$$

$$TV = \int_0^\infty |u(t+1) - u(t)| = \sum_{k=1}^{K-1}|u(k) - u(k-1)|, \tag{4}$$

where K is the total number of sampling intervals in the experiment. Again, note that Δt in Eq. (3) represents the monitoring sampling time, assumed to be smaller than the one implicit in Eq. (2). A critical issue in optimization-based PID tuning is the type of process disturbance considered for the evaluation of these performance metrics. In this work, we include in the same experiment both step setpoint changes as well as step input disturbances for each loop. To reduce the number of

performance metrics, and because the use of setpoint weighting enables a controller structure with two degrees of freedom, no distinction between servo and regulatory performance is explicitly made.

The tuning goal is to achieve a balance of tracking performance and smooth control by simultaneously minimizing both IAE and TV for all loops. Because these are usually conflicting objectives, due to variable coupling between loops and the trade-off between performance and control effort for each loop, the selection of controller parameters is formulated as a multi-objective optimization problem solved using the weighted sum method (Marler and Arora, 2010):

$$J(x) = \sum_{i=1}^{n} \left[\alpha_i IAE_i(x) + \beta_i TV_i(x)\right], \tag{5}$$

where J is the objective function, x is the vector of controller parameters, α and β are positive weights that regulate the trade-offs between setpoint tracking performance and control effort for each loop i, and relative importance of each loop. By selecting different values for α and β for each loop, the controllers can be tuned to reflect desired performance. For instance, detuning of less relevant loops is achieved by choosing lower values or even zero weighting, for α. To correctly assign preferences, the objectives should be standardized to similar magnitudes (Marler and Arora, 2010). This is done by first transforming them to their dimensionless counterparts, Normalized IAE, $NIAE$, and Normalized TV, NTV:

$$NIAE = \sum_{k=0}^{K} \frac{|y_{sp}(k) - y(k)|}{y_{sp}(k)} \Delta t \tag{6}$$

$$NTV = \sum_{k=1}^{K-1} \frac{|u(k) - u(k-1)|}{u_{max} - u_{min}}. \tag{7}$$

Then, the objectives are also divided by their median value, med, inside the optimization domain, leading to the normalized objective function, J:

$$J(x) = \sum_{i=1}^{n} \left[\frac{\alpha_i}{\text{med}(NIAE_i)} NIAE_i(x) + \frac{\beta_i}{\text{med}(NTV_i)} NTV_i(x)\right] \tag{8}$$

The median value is calculated based on a initial set of space-filling samples used for BO, as will be described in the next section.

2.3. Optimization-based PID controller tuning

The optimization-based tuning problem of n decentralized PID controllers is summarized as the following black-box multi-objective optimization problem:

$$\min_{x \in \mathbb{X}} \; J(x) \tag{9a}$$

$$\text{s.t. } x = [\Delta_{ti}, K_{Pi}, K_{Ii}, b_i, \dots, \Delta_{tn}, K_{Pn}, K_{In}, b_n], i = 1, \dots, n$$

$$K_P, K_I \in \mathbb{R}^{2n}, \; \Delta_t \in \mathbb{R}^{+n}, \; b \in [0,1]^n \tag{9b}$$

$$\mathbb{X} = \{x \mid x_i^{lb} \le x_i \le x_i^{ub}, i = 1, \dots, n\} \tag{9c}$$

We are interested in finding the global (or a "good enough") optimum of J. Note that, even when a model is available, there is usually no closed-form expression for the objective function or its gradients in the case of non-linear multivariable processes or non-linear controllers (with saturation). Thus, the objective function can be thought of as a black-box whose response surface can be inferred using closed-loop experiments. If these are done on the real process, each evaluation is expensive, which means that sample efficient approaches are required, rendering standard Derivative Free Optimization (DFO) methods impracticable (Sumana and Venkateswarlu, 2010; Xue et al., 2010; Dittmar et al., 2012). In this work, we propose to solve the optimization problem (9) with BO, which is an efficient method for noisy and expensive black-box DFO. In the proposed framework, controller parameters are optimized based on iterative closed-loop experiments, as displayed in Fig. 2.

The proposed framework has the flexibility to consider other arbitrarily specified performance metrics as well as other disturbances for the evaluation of J. Constraints on the desired closed-loop response related to safety or quality measures can also be readily considered (Schillinger et al., 2017; Sorourifar et al., 2021; Khosravi et al., 2022b). Instead of using the weighted sum method, multi-objective BO schemes that approximate the Pareto front can also be employed (Makrygiorgos et al., 2022). However, under the considered performed metrics for multi-loop PID tuning, this leads to a high-dimensional ($2n$) Pareto front, which is hard to visualize and approximate. Finally, the vector of decision variables x could be easily extended to include derivative gains, and, for a more complete automatic tuning procedure, parameters related to measurement noise filtering or anti-windup measures.

3. Proposed methodology

The main problem when using a data-driven method such as BO to solve the optimization-based tuning problem (9) is the definition of the upper, x^{ub}, and lower, x^{lb}, bounds on the controller parameters without any prior knowledge. A simple approach for an operating process would be to simply detune the parameters of each controller by constant factors. However, due to the arbitrary nature of this approach, the risk of closed-loop instability increases significantly.

In this section, we propose a systematic methodology that combines sequential loop closing, system identification and model-based tuning relations to define a relevant optimization domain. For generality, we assume no prior knowledge regarding the range of controller parameters or that a MIMO process model is available. However, we assume that the controller pairings are already correctly determined using suitable methods and that some form of qualitative knowledge regarding loop speed is available, which is usually the case for operating processes. A schematic representation of the proposed methodology is presented in Fig. 3.

The proposed methodology can be seen as a data-driven and automatic method to refine the controllers parameters obtained using a sequential tuning for desired performance. In this approach, closed-loop stability is approximately assured by constraining the controller parameters to relevant ranges, around initially tuned controllers that lead to stable closed-loop responses.

3.1. Sequential system identification and design space definition

To describe the open-loop behavior between y and u for each loop, we consider two types of Single-Input Single Output (SISO) low-order models which generally adopted in the design of PID controllers in industry, the First Order Plus Time Delay (FOPTD) model,

$$G(s) = \frac{K_p}{\tau s + 1} e^{-\theta s}, \tag{10}$$

and the Integrating plus Time Delay model (ITD) model:

$$G(s) = \frac{K_v}{s} e^{-\theta s}, \tag{11}$$

where, K_v is the process slope, K_p its static gain, τ its time constant, and θ its delay. For a multi-loop feedback system, the open-loop transfer function between a process input and output with all the other loops closed actually depends on all transfer function elements and other controllers, through a hidden feedback path (Shen and Yu, 1994; Vu and Lee, 2010). This means that an independent open-loop system identification approach, with all loops on manual mode, is not able to capture the process interactions, especially under closed-loop.

A simple and straightforward approach to take into account process interactions consists of using a sequential approach for system identification (Shen and Yu, 1994), where n process transfer functions for n decentralized control loops are identified by closing each loop sequentially. As represented in Fig. 3, the sequential approach starts

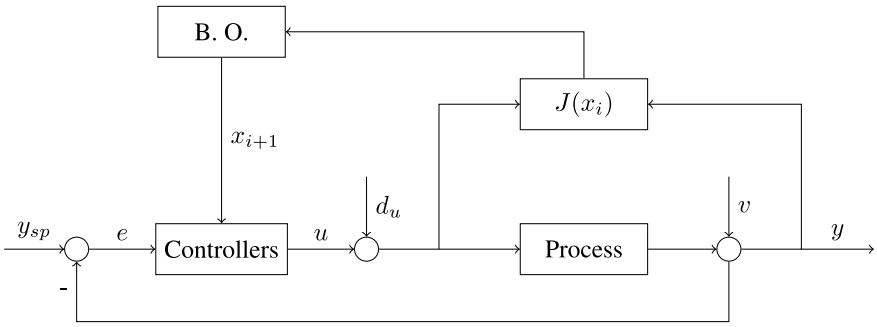

Fig. 2. PID controller tuning using Bayesian optimization (B.O.), where i denotes a closed-loop experiment with parameters x_i.

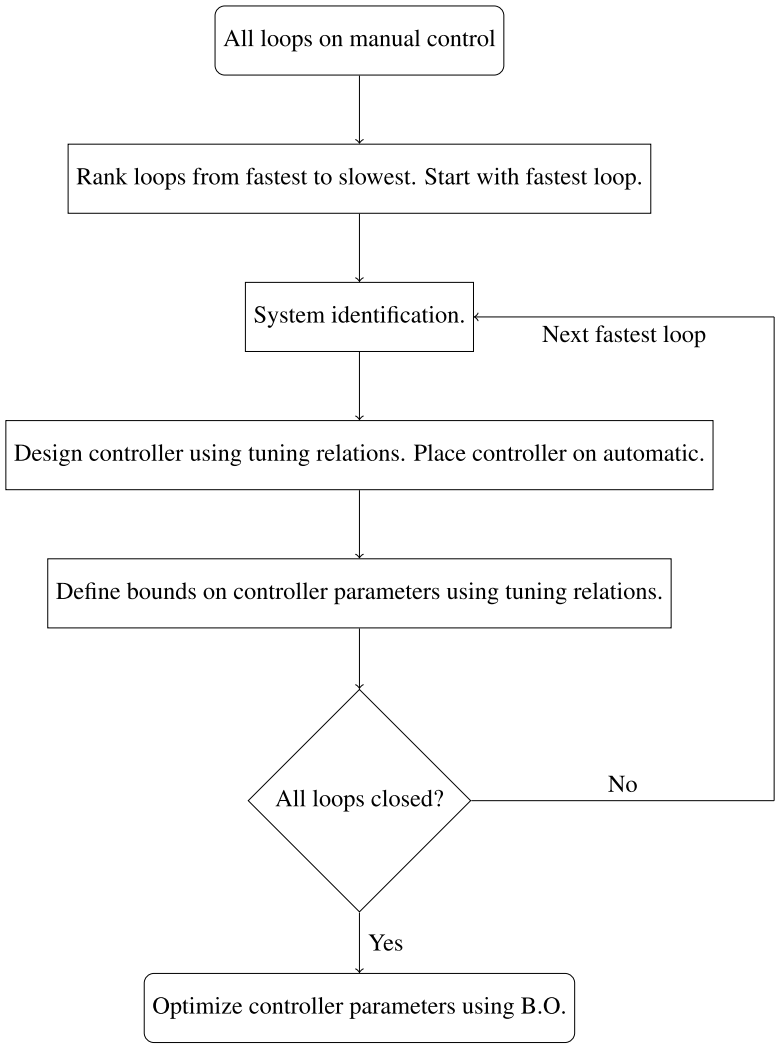

Fig. 3. Flowchart of the proposed methodology.

with all loops open, under manual control. System identification is performed on the fastest loop and the controller is tuned based on the identified model and placed on automatic control. Then, the second fastest loop is identified, tuned and closed. This procedure is repeated until the last, slowest, loop is identified, tuned and closed. The sequential approach has the advantages that simpler SISO identification methods could be used and only n (instead of $n \times n$) transfer functions needs to be identified.

In this work, we opted for efficient methods of system identification, namely relay feedback tests, where the system is under feedback control by means of a relay (on–off controller), which results in sustained oscillations of the controlled variable. Based on the characteristics of these oscillations, the parameters of a simple FOPTD or ITD process model can be estimated (Liu et al., 2013). We employ the method developed by Berner et al. (2016), due to its simplicity, no prior knowledge required, software availability and previous experimental validation of the methodology (Berner et al., 2016, 2018). This method uses an asymmetric relay test to excite the system in order to estimate the process gain and normalized time delay. The latter is used to determine whether an ITD or FOPTD model is estimated. More details regarding the method can be found in Berner et al. (2016). We would like to point out that other methods could have been used to excite

the system, such as open-loop step tests or even closed-loop step tests, under P or PI control (Liu et al., 2013).

During the sequential system identification, the controllers are tuned using the SIMC relations (Skogestad, 2003), which give the following controller parameters for a FOPTD process (after conversion to the expanded form):

$$K_P = \frac{1}{K_p}\frac{\tau}{\tau_C+\theta}, \quad K_I = \frac{1}{K_p}\frac{\tau}{\tau_C+\theta}\frac{1}{\min\{\tau,\ 4(\tau_C+\theta)\}}. \tag{12}$$

For an ITD process:

$$K_P = \frac{1}{K_v}\frac{1}{\tau_C+\theta}, \quad K_I = \frac{1}{K_v}\frac{1}{\tau_C+\theta}\frac{1}{4(\tau_C+\theta)}. \tag{13}$$

The desired closed-loop time constant, τ_C, can be selected based on the value of θ, usually with $\tau_C = \theta$ (Skogestad, 2003; Grimholt and Skogestad, 2018). In this work, we propose a systematic method to define bounds on the controller parameters for each loop based on these tuning relations and on the identified transfer function process model. The bounds on Δt are defined based on available heuristics, such as (Åström and Wittenmark, 2011):

$$\Delta t^{lb} = 0.01\tau, \quad \Delta t^{ub} = 0.05\tau. \tag{14}$$

Other guidelines can be found in Isermann (1989) and Åström and Wittenmark (2011). For ITD processes, the bounds can be defined arbitrarily or using heuristics. Because the digital controller sampler introduces an effective closed-loop time delay which affects stability, the bounds of the controller gains should depend on Δt. This dependence can be accounted by approximating the closed-loop sampling delay as $0.5\Delta t$ (Ogunnaike and Ray, 1994), and obtaining a range of values for the modified time delay, θ':

$$\theta'_{min} = \theta + 0.5\Delta t^{lb}, \quad \theta'_{max} = \theta + 0.5\Delta t^{ub}. \tag{15}$$

The bounds on the controller gains can be obtained by simply varying τ_C between a minimum, $\tau_{C,min}$, and a maximum, $\tau_{C,max}$, value. These can be based on the θ', with $\tau_{C,min} = 0.5\theta'_{min}$ and $\tau_{C,max} = 1.5\theta'_{max}$. For a FOPTD process, the bounds are given by:

$$K_P{}^{ub} = \frac{1}{K_p}\frac{\tau}{\tau_{C,min}+\theta'_{min}} = \frac{\tau}{K_p 1.5\theta'_{min}} \tag{16}$$

$$K_P{}^{lb} = \frac{1}{K_p}\frac{\tau}{\tau_{C,max}+\theta'_{max}} = \frac{\tau}{K_p 2.5\theta'_{max}} \tag{17}$$

$$K_I{}^{ub} = \frac{1}{K_p}\frac{\tau}{\tau_{C,min}+\theta'_{min}}\frac{1}{\min\{\tau,\ 4(\tau_{C,min}+\theta'_{min})\}} = \frac{\tau}{K_p 1.5\theta'_{min}}\frac{1}{\min\{\tau,\ 6\theta'_{min}\}} \tag{18}$$

$$K_I{}^{lb} = \frac{1}{K_p}\frac{\tau}{\tau_{C,max}+\theta'_{max}}\frac{1}{\min\{\tau,\ 4(\tau_{C,max}+\theta'_{max})\}} = \frac{\tau}{K_p 2.5\theta'_{max}}\frac{1}{\min\{\tau,\ 10\theta'_{max}\}}. \tag{19}$$

Similarly, for an ITD process:

$$K_P{}^{ub} = \frac{1}{K_v}\frac{1}{\tau_{C,min}+\theta'_{min}} = \frac{1}{K_v 1.5\theta'_{min}} \tag{20}$$

$$K_P{}^{lb} = \frac{1}{K_v}\frac{1}{\tau_{C,max}+\theta'_{max}} = \frac{1}{K_v 2.5\theta'_{max}} \tag{21}$$

$$K_I{}^{ub} = \frac{1}{K_v}\frac{1}{\tau_{C,min}+\theta'_{min}}\frac{1}{4(\tau_{C,min}+\theta'_{min})} = \frac{1}{K_v 9{\theta'_{min}}^2} \tag{22}$$

$$K_I{}^{lb} = \frac{1}{K_v}\frac{1}{\tau_{C,max}+\theta'_{max}}\frac{1}{4(\tau_{C,max}+\theta'_{max})} = \frac{1}{K_v 25{\theta'_{max}}^2}. \tag{23}$$

For processes with negative gain, the upper and lower bounds are switched accordingly.

3.2. Bayesian optimization

BO is a class of surrogate model-based Derivative Free Optimization (DFO) algorithms for expensive and noisy black-box optimization problems (Brochu et al., 2010; Shahriari et al., 2016; Greenhill et al., 2020). The two main components of BO are: (*i*) a probabilistic surrogate model used to approximate the underlying objective function, most commonly a Gaussian Process (GP) and (*ii*) an auxiliary, or acquisition, function that is optimized to determine which experiments to perform in an iterative search for the global optimum.

3.2.1. Gaussian process regression

GPR is a non-parametric probabilistic method where the underlying function is modeled as a Gaussian Process (GP), formally defined as a collection of random variables, with a joint Gaussian distribution (Rasmussen and Williams, 2006). Using GPR, the objective function, $J(x)$, is approximated as a GP, which is completely specified by its mean, $m(x)$, and covariance function, $k(x,x')$:

$$J(x) \sim \mathcal{GP}(m(x), k(x,x')), \tag{24}$$

where x and x' are D-by-1 vectors, where D=4n is the number of input variables, in this case, the controller parameters defined in the previous section. Both the mean and covariance functions can encode prior knowledge regarding J. Without loss of generality, we consider a zero mean GP, with $m(x) = 0$. The covariance function, or kernel, $k(x,x')$ may encode prior knowledge about the smoothness of J and how the objective function at a given point is influenced by other points in the neighborhood. While several types of kernels are available (Rasmussen and Williams, 2006), in this work, we choose the Matérn 5/2 kernel, which is adequate for modeling functions that are twice differentiable, and not overly smooth (Snoek et al., 2012). We further consider the anisotropic version of this kernel, with Automatic Relevance Determination (ARD):

$$k(x,x') = {\sigma_f}^2\left(1+\sqrt{5r^2(x,x')}+\frac{5}{3}r^2(x,x')\right)\exp\left(-\sqrt{5r^2(x,x')}\right), \tag{25}$$

with,

$$r^2(x,x') = (x-x')^T M (x-x'). \tag{26}$$

σ_f is the signal standard deviation and M is a diagonal matrix, $M = \mathrm{diag}(l)^{-2}$, where l is a vector of positive length-scale parameters $l = [l_1,\dots,l_\mathrm{D}]$. These length-scales reflect the distance traveled along the dimension $d = 1,\dots,\mathrm{D}$ of the input space before the underlying function value changes significantly. Larger values of l_d imply that the objective function changes more slowly with regards to input variable d, meaning that this input is less relevant to the covariance, effectively implementing ARD (Rasmussen and Williams, 2006).

Since the system is subject to measurement noise, we have only access to noisy evaluations, f, of the objective function J:

$$f(x) = J(x) + \varepsilon, \tag{27}$$

where, ε is the observation noise, assumed to be an independent and identically distributed random variable that follows a zero-mean Gaussian distribution, $\varepsilon \sim \mathcal{N}(0,{\sigma_n}^2)$. Let us consider that t previously obtained noisy measurements of the objective function are available, $f_{1:t} = \{f_1,\dots,f_t\}$, evaluated at the corresponding controller parameters $x_{1:t} = \{x_1,\dots,x_t\}$, which we denote as the training dataset, $D_t = \{x_i, f_i\}_{i=1}^t$. With the assumption of Gaussian noise, and under the GP prior, the joint distribution of the noisy training outputs, $f_{1:t}$, and the testing output, $J(x)$, evaluated at any new controller parameters, x, is also Gaussian:

$$\begin{bmatrix} f_{1:t} \\ J(x) \end{bmatrix} \sim \mathcal{N}\left(0,\ \begin{bmatrix} \mathrm{K}+{\sigma_n}^2 I & \mathrm{k} \\ \mathrm{k}^T & k(x,x) \end{bmatrix}\right), \tag{28}$$

where I is the identity matrix with size t, k is the vector of covariances between x and $x_{1:t}$, and K is the matrix of covariances evaluated at all pairs of training inputs, $x_{1:t}$:

$$\mathrm{k} = [k(x,x_1),\dots,k(x,x_t)] \tag{29}$$

$$\mathrm{K} = \begin{bmatrix} k(x_1,x_1) & \cdots & k(x_1,x_t) \\ \vdots & \ddots & \vdots \\ k(x_t,x_1) & \cdots & k(x_t,x_t) \end{bmatrix}. \tag{30}$$

The posterior distribution of $J(x)$ at any x, conditioned on the data D_t, can then be computed analytically:

$$J(x) \mid D_t \sim \mathcal{N}(\mu_t(x), {\sigma_t}^2(x)), \tag{31}$$

where the predictive mean, $\mu_t(x)$, and variance, ${\sigma_t}^2(x)$, are given by:

$$\mu_t(x) = \mathrm{k}^T[\mathrm{K} + {\sigma_n}^2 I]^{-1} f_{1:t} \tag{32}$$

$${\sigma_t}^2(x) = k(x,x) - \mathrm{k}^T[\mathrm{K} + {\sigma_n}^2 I]^{-1}\mathrm{k}. \tag{33}$$

The mean, $\mu_t(x)$, provides the best guess of $J(x)$ and the variance, ${\sigma_t}^2(x)$, offers a measure of the associated prediction uncertainty. The GP model is parameterized by several, usually unknown, hyper-parameters, such as signal and noise variances as well as length-scales. Although these can be fixed a priori, in this work, the hyper-parameters are calibrated to the available data by maximizing the marginal log-likelihood (Rasmussen and Williams, 2006):

$$\log p(f_{1:t} \mid x_{1:t}, \theta) = -\frac{1}{2} {f_{1:t}}^T[\mathrm{K} + {\sigma_n}^2 I]^{-1} f_{1:t} - \frac{1}{2}\log|\mathrm{K} + {\sigma_n}^2 I| - \frac{t}{2}\log 2\pi, \tag{34}$$

where $\theta = [\sigma_f, \sigma_n, l_1, \ldots, l_D]$ is the vector of GP hyper-parameters. For further details on GPR, the reader is referred to Rasmussen and Williams (2006).

3.2.2. Bayesian optimization

The BO framework uses the predictive mean and variance of the GP model to formulate an acquisition function that trades off exploitation, testing promising controller parameters given the current knowledge, with exploration, sampling unexplored regions of the design space. BO was implemented using the `bayesopt` function in the Statistics and Machine Learning Toolbox™ in Matlab 2020b. A short general description of the BO algorithm is presented in Algorithm 1.

Algorithm 1 Bayesian Optimization Algorithm

Require: Domain $\mathbb{X}$, initial data D_0 and maximum number of experiments, N_{max}

for $t = 1, \ldots, N_{max}$ **do**

Use current dataset D_{t-1} to train the GP model and calculate the posterior distribution via (31)

Maximize the acquisition function, α, to find x_t : $x_t = \arg\max_{x \in \mathbb{X}} \alpha(x; D_{t-1})$

Evaluate the objective function at x_t to obtain f_t

Augment the dataset with (x_t, f_t): $D_t = \{D_{t-1}, (x_t, f_t)\}$

end for

Return: Value x_{t^*} that led to the lowest observed objective function value, with $t^* = \arg\min_{t \in \{1,\ldots,N_{max}\}} f_t$

The BO algorithm can be initiated without any past observations, but it is usually more efficient to calibrate the GP model hyper-parameters and calculate a prior by first collecting an initial set of observations $D_0 = \{x_i, f_i\}_{i=1}^{t_0}$. To adequately cover the optimization domain, $x_{1:t_0}$ are obtained with a space-filling Latin Hypercube Sampling (LHS) design (McKay et al., 1979). In this work we set the initial sample size to be half of the total experiment budget, with $t_0 = 0.5N_{max}$. The algorithm then proceeds with a loop where in each iteration, given the available data, the hyper-parameters of the GP model are obtained by maximizing Eq. (34), and the posterior distribution in Eq. (31) is recalculated. The next controller parameters to evaluate, x_t, are found by maximizing an auxiliary utility, or acquisition, function, α, based on current predictive mean and variance in Eqs. (32) and (33), respectively:

$$x_t = \arg\max_{x \in \mathbb{X}} \alpha(x; D_t). \tag{35}$$

Since this auxiliary function is cheaper to evaluate than the original objective function, $J(x)$, common DFO or gradient-based algorithms can be used for its optimization. In this work, we consider the Expected Improvement (EI) acquisition function (Jones et al., 1998), which calculates the expectation of the improvement brought by a given x, over the best current solution, or incumbent, η. EI can be expressed analytically as:

$$\alpha(x; D_t) = \mathbb{E}[\eta - J(x)] = (\eta - \mu_t(x))\Phi\left(\frac{\eta - \mu_t(x)}{\sigma_t(x)}\right) + \sigma_t(x)\phi\left(\frac{\eta - \mu_t(x)}{\sigma_t(x)}\right), \tag{36}$$

where Φ and ϕ are the standard normal cumulative distribution and probability density functions, respectively. By definition, this acquisition function should only consider positive values: $\alpha(x; D_t) = \mathbb{E}[\max\{0, \eta - J(x)\}]$. Considering that the objective function is noisy, the incumbent η, is defined as the value that minimizes the posterior mean of the GP model at the current iteration, $\eta = \min_x \mu_t(x)$. Other types of acquisition functions are presented in Shahriari et al. (2016). After having found controller parameters x_t, a closed-loop experiment is performed to obtain f_t and the dataset is updated. This procedure is iterated until a pre-determined stopping criterion is met, which in this work is considered as the maximum number of allowed experiments, N_{max} is achieved. This value can be determined, for instance, based on the total time allocated for controller tuning and the duration of each closed-loop experiment. Finally, after the stopping criteria is met, the optimum controller parameters (given the available budget) are considered as the ones that led to the best observed objective function value. Alternatively, in an highly noisy environment, the best controller parameters could be considered to be those that minimize the final GP model posterior mean, $\mu_{N_{max}}(x)$, or confidence bound $\mu_{N_{max}}(x) - \sigma_{N_{max}}(x)$.

4. Simulated case study: Non-linear evaporator

In this section, we compare the proposed methodology with sequential loop closing using SIMC relations, in order to assess how much the initial performance can be improved with BO. The simulations are performed using Simulink® and the relay identification implementation available in https://gitlab.control.lth.se/Josefin/AutotunersMatlab. As a case study we consider a non-linear evaporator process proposed by Lee et al. (1989), used as a benchmark for multivariable process control studies (Dittmar, 2015). A schematic representation of this process is shown in Fig. 4.

Assuming fast dynamics for the energy balances, the process model is described by the following set of differential and algebraic equations:

$$\frac{dL_2}{dt} = \frac{F_1 - F_2 - F_4}{\rho A} \tag{37}$$

$$\frac{dX_2}{dt} = \frac{F_1 X_1 - F_2 X_2}{M} \tag{38}$$

$$\frac{dP_2}{dt} = \frac{F_4 - F_5}{C} \tag{39}$$

$$F_4 = \frac{Q_{100} - F_1 C_p (T_2 - T_1)}{\lambda} \tag{40}$$

$$F_5 = \frac{Q_{200}}{\lambda} \tag{41}$$

$$Q_{200} = \frac{U A_2 (T_3 - T_{200})}{1 + \frac{U A_2}{2 C_p F_{200}}} \tag{42}$$

$$Q_{100} = 0.16(F_1 + F_3)(T_{100} - T_2) \tag{43}$$

$$T_{100} = 0.1538 P_{100} + 90 \tag{44}$$

$$T_2 = 0.5616 P_2 + 0.3126 X_2 + 48.43 \tag{45}$$

$$T_3 = 0.507 P_2 + 55. \tag{46}$$

This process has 3 controlled variables, the separator level, L_2, solute exit concentration, X_2, and evaporator pressure, P_2 and 3 manipulated

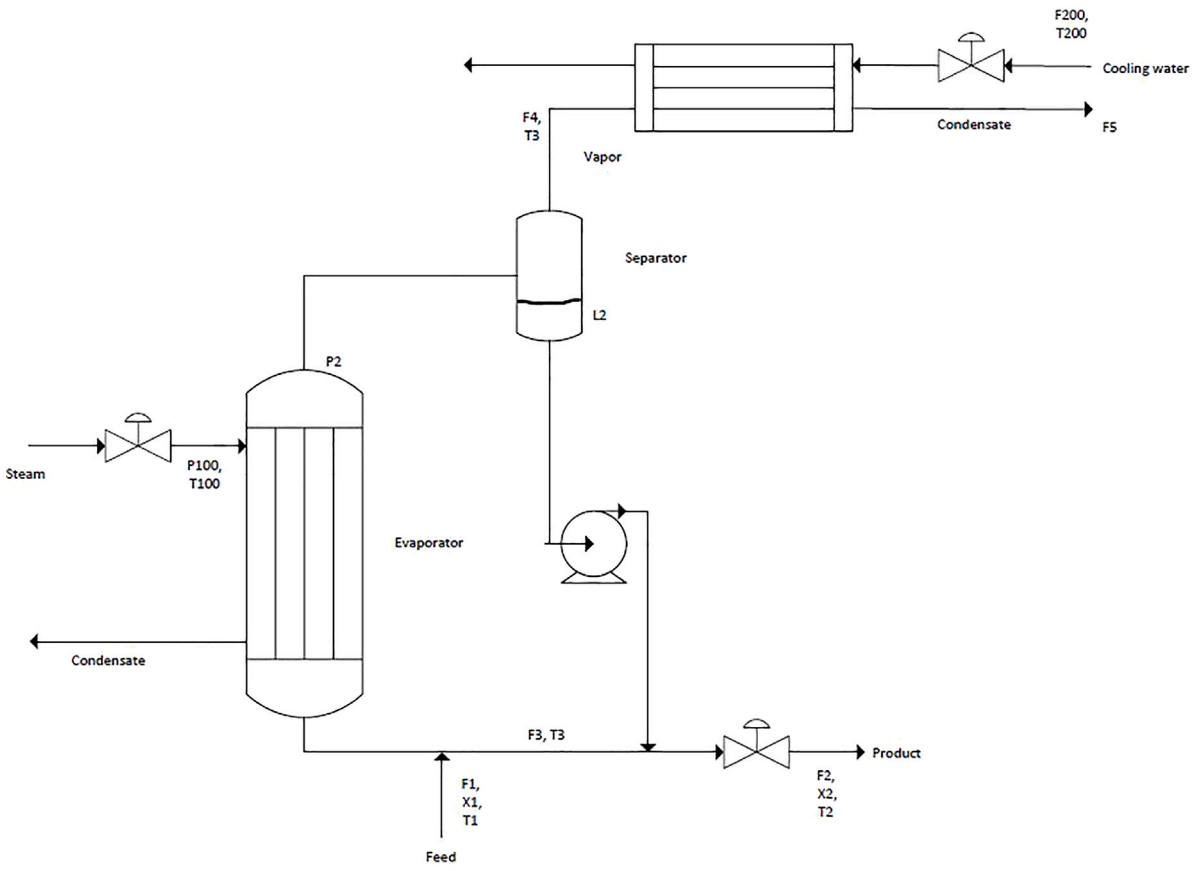

Fig. 4. Evaporator process.

Table 1
Evaporator model parameters and nominal operating conditions.

Parameter	Value	Variable	Value
ρA	20 kg/m	L_2	1 m
M	20 kg	P_2	50.5 kPa
C	4 kg/kPa	X_2	25%
C_p	0.07 kW min/(kg K)	F_2	2 kg/min
λ	38.5 kW min/kg	P_{100}	194.7 kPa
UA_2	6.84 kW/K	F_{200}	208 kg/min
F_1	10 kg/min	X_1	5%
T_1	$40°C$	F_3	50 kg/min
T_{200}	$25°C$		

variables, product flowrate, F_2, steam supply pressure, P_{100}, and the condenser cooling water flowrate, F_{200}. The disturbance variables are inlet flowrate, F_1, solute concentration, X_1, and temperature, T_1, the recycle flowrate, F_3, and cooling water supply temperature, T_{200}. The model parameters and nominal operating conditions are displayed on Table 1.

We consider the same controller structure as in Lee et al. (1989): L_2 is controlled by manipulating F_2 (loop 1), P_2 is controlled by manipulating F_{200} (loop 2) and X_2 is controlled by manipulating P_{100} (loop 3). In addition to being non-linear, the system is highly interactive under nominal operating conditions (Lee et al., 1989), providing a challenging benchmark to test the proposed methodology for decentralized PID controller tuning. The original model by Lee et al. (1989) did not consider time delays, which are however present in most industrial chemical processes. Therefore, transport delays of 1 and 2 min were added to L_2 and P_2, respectively, while X_2 has a measurement delay of 10 min. Also, the measurements of L_2, P_2 and X_2 are subjected to additive zero-mean Gaussian noise with standard deviations of 0.01 m, 0.15 kPa and 0.25%.

Noise-free open loop step tests show that each loop displays distinct dynamic behavior: $L_2 - F_2$ is integrating, and, with L_2 under automatic control, $P_2 - F_{200}$ is apparently of first order while $X_2 - P_{100}$ is apparently of second order with a negative zero (Dittmar, 2015). However, in the tuning procedure, we only consider the relative loop speed to be known in advance: loop 1 is closed first (since it is integrating), followed by loop 2, which is faster than loop 3. Following this sequence, the sequential system identification leads to the following models:

$$G_1(s) = \frac{-0.076}{s} e^{-1.72s} \tag{47}$$

$$G_{2,cl}(s) = \frac{-0,003}{s} e^{-2.62s} \tag{48}$$

$$G_{3,cl}(s) = \frac{0.327}{4.79s + 1} e^{-13.77s}, \tag{49}$$

where $G_{2,cl}$ represents the transfer function between $P_2 - F_{200}$ after closing loop 1, and $G_{3,cl}$ the transfer function between $X_2 - P_{100}$ with loops 1 and 2 closed. Fig. 5 shows the comparison between the true process response and that of the identified model subjected to the same relay test inputs.

It can be observed that the response of models G_1 and $G_{3,cl}(s)$ are close to that of the true process but the response of $G_{2,cl}(s)$ presents a large mismatch. In practice, the system identification test could be repeated to obtain a better model. However, here, we choose to keep the model with mismatch to display the advantages of the proposed methodology even when there is a mismatch between the identified model and the true process. Since an ITD model was identified for G_1 and G_2, we selected the values of Δt_1 and Δt_2 considering the recommend sampling times of 10 and 5 s (0.17 and 0.083 min) for level and pressure loops, respectively (Isermann, 1989). Using the proposed methodology leads to the optimization domain displayed on Table 2.

A total budget of 20 closed-loop experiments is considered to optimize the controller parameters. In each experiment, we consider +10%

Table 2
Domain of controller parameters and manipulated variables for evaporator process (loop 1: $L_2 - F_2$, loop 2: $P_2 - F_{200}$, loop 3: $X_2 - P_{100}$).

Variable	Range	Variable	Range	Variable	Range
Δt_1	[0.1, 0.5]	Δt_2	[0.05, 0.5]	Δt_3	[0.048, 0.240]
K_{P1}	[−4.730, −2.561]	K_{P2}	[−75.461, −41.779]	K_{P3}	[0.420, 0.706]
K_{I1}	[−0.426, −0.125]	K_{I2}	[−4.680, −1.435]	K_{I3}	[0.088, 0.147]
b_1	[0, 1]	b_2	[0, 1]	b_3	[0, 1]
F_2	[0, 5]	F_{200}	[0, 400]	P_{100}	[0, 400]

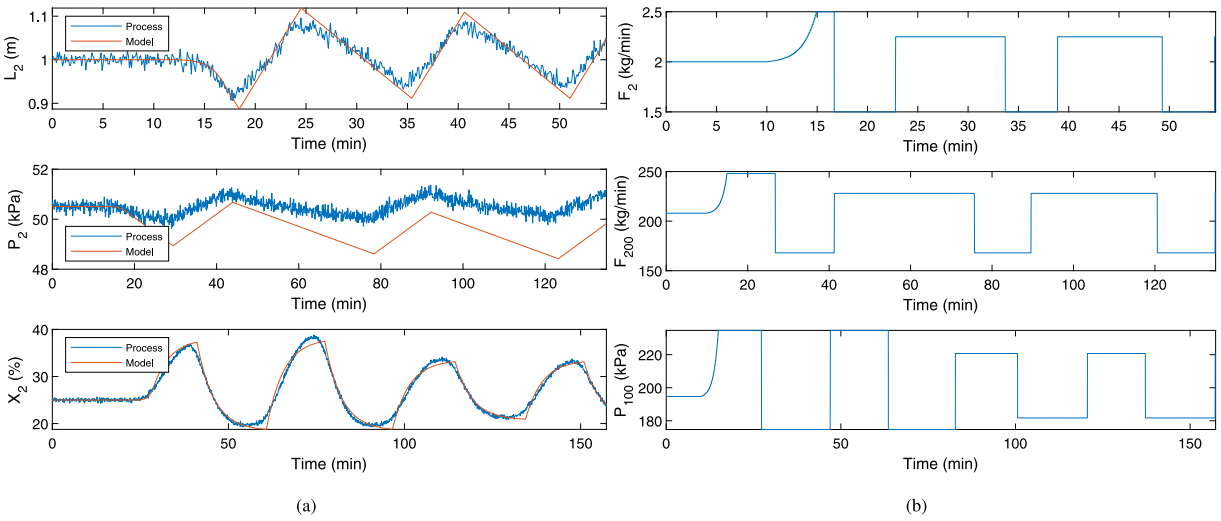

Fig. 5. Relay system identification results: (a) comparison between true process and identified model output under the relay test, (b) manipulated variables.

setpoint changes in L_2, P_2 and X_2 at 0, 100 and 250 min, respectively, as well as +10% step input disturbances in F_2, F_{200} and P_{100} at 50, 150 and 350 min, respectively. In this example, we define the objective weights, α and β, to reflect the following desired control performance specifications, ranked from most to least important:

- X_2 is a quality variable that must be tightly controlled.
- F_2 must have smooth variations, to enable some reduction of fluctuations in production rate.
- The remaining performance metrics have equal weighting and importance.

Based on these criteria, $NIAE_3$ is given the highest weight value ($\alpha_3 = 1000$), followed by NTV_1 ($\beta_1 = 10$). All the other criteria are given a value of 1. $NIAE$ is calculated considering a monitoring sampling time of 2 s (0.03 min) for all loops. The convergence of BO is impacted by several sources of variability, including measurement noise, the inner optimization algorithm and the initial sample proposed by LHS. To assess the impact of these factors, we perform 30 independent runs of BO with different measurement noise realizations and initial samples by varying the random number generator seed between iterations. Because the purpose of this analysis is not to assess the impact of noise on system identification, the optimization domain remains constant, as defined in Table 2, using the models G_1, G_2 and G_3. The results of this analysis are displayed on Fig. 6.

In Fig. 6, we can observe that even in the presence of measurement noise, and despite the inner stochastic optimization problem, the performance of BO is consistently improved with additional experiments. To illustrate the differences in closed-loop performance obtained due to the noise variability during BO, the controller parameters corresponding to the best and worst final values of the 30 runs are selected for comparison with the sequential SIMC approach. These are displayed on Table 3, together with the different performance metrics, including the Relative Improvement (RI) over the sequential SIMC method. The closed-loop responses are shown in Fig. 7.

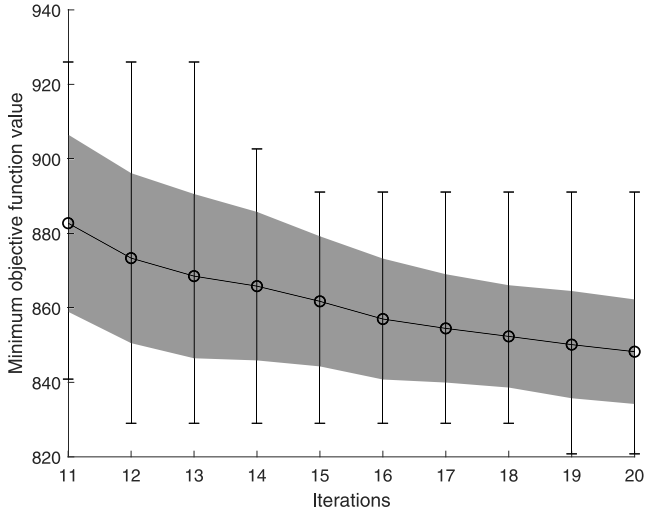

Fig. 6. Best objective function values during the BO procedure. Circles represent the mean values, the shaded region +/− one standard deviation and the vertical lines the minimum and maximum values of the 30 runs.

In can be verified that, in both cases, BO selects higher values of Δt for all loops, possibly to reduce the effect of high frequency measurement noise on the controller output. This leads to lower values of NTV when compared to SIMC. For loop 2, in particular, SIMC leads to a value of NTV that is roughly 3 times as large as the one obtained with BO. The most apparent differences in the controllers parameters found in the best and worst BO trials take place in loops 1 and 2. For these loops, the behavior of the controlled variable obtained with the worst parameters of BO is similar to the one obtained with SIMC, yet with a lower NTV. In spite of differences in performance of BO, the normalized objective function value and the two most

Table 3
Controller parameters and performance metrics for evaporator process (C_1: $L_2 - F_2$, C_2: $P_2 - F_{200}$, C_3: $X_2 - P_{100}$).

Controller	Method	Δt	K_P	K_I	b	$NIAE$	NTV	$J_{norm} \times 10^3$	RI (%)
C_1	Sequential SIMC ($\tau_C = \theta$)	0.17	−3.65	−0.25	1	8.21	22.99	1.05	–
	BO (best)	0.37	−4.55	−0.26	0.57	6.40	12.59	0.82	22
	BO (worst)	0.45	−3.54	−0.39	0.07	7.69	8.44	0.84	20
C_2	Sequential SIMC ($\tau_C = \theta$)	0.08	−57.12	−2.68	1	1.84	131.04	1.04	–
	BO (best)	0.38	−72.39	−1.45	0.23	4.78	38.22	0.82	22
	BO (worst)	0.32	−74.56	−4.24	0.75	1.97	44.08	0.84	20
C_3	Sequential SIMC ($\tau_C = \theta$)	0.05	0.53	0.11	1	19.37	3.63	1.04	–
	BO (best)	0.23	0.67	0.15	0.96	15.16	1.04	0.82	22
	BO (worst)	0.13	0.63	0.15	0.88	15.74	1.61	0.84	20

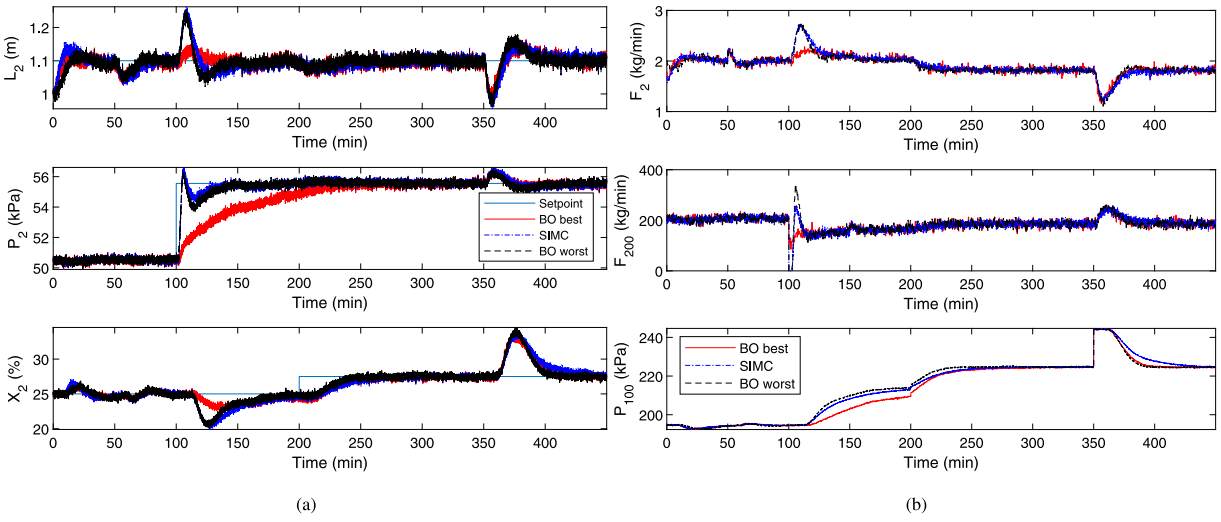

Fig. 7. Closed-loop responses of the evaporator process subject to different setpoint changes and disturbances: (a) controlled variables, (b) manipulated variables.

important metrics, $NIAE_3$ and NTV_2, are similar in both the worst and best cases, which represent an increase of 22% and 20% in relative performance, respectively, when compared with the SIMC method. It is important to note that both these trials represent extreme cases, since they correspond to the maximum and minimum values. Therefore, we argue that closed-loop performance obtained with controllers tuned by BO remains consistent despite measurement noise and the stochastic nature of the BO algorithm. Although SIMC leads to overall satisfactory closed-loop behavior, tuning the controllers to achieve these specific performance criteria would be hard to do using trial and error detuning in an interactive process such as this one. Moreover, this 20% increase in performance is obtained with only 20 experiments and considering the optimization of 12 simultaneous controller parameters, which indicates the data-efficiency of BO.

5. Conclusions

In this work, we presented a novel and systematic methodology for multi-loop PID controller tuning combining BO with sequential loop closing and model-based SIMC tuning relations. After defining the optimization domain, the parameters of all the multi-loop controllers are simultaneously optimized through on-line closed-loop experiments with BO, based on performance specifications.

Simulation results show that the performance obtained using a sequential approach and SIMC tuning relations, although satisfactory, can be significantly improved with our proposed methodology, especially considering specific trade-offs in the desired performance of each decentralized controller. This is done without iterative re-opening of the control loops or development of a MIMO process model, which represent advantages compared with using a purely sequential tuning approach or off-line optimization-based tuning, respectively.

There are several interesting research directions that could be considered for further improving the applicability of BO for automatic multi-loop PID tuning. BO is well-known to have slower convergence for high dimensional problems, which can arise when tuning a larger number of controllers. Possible solutions to improve scalability of BO include methods which exploit an intrinsic lower dimensionality of the problem (Shahriari et al., 2016), or adapt the optimization domain in a dynamic fashion (Fröhlich et al., 2019). BO requires several on-line closed-loop experiments, which increases the autotuning cost, time, and the possibility of occurring non-stationary external disturbances or process variations. Therefore, future extensions to this work include the use of multi-fidelity BO schemes (Marco et al., 2017) to lower the on-line experimental cost and time of BO. Alternatively, contextual BO approaches (Fiducioso et al., 2019) could also be used, resulting in adaptive multi-loop PID control. Finally, it would also be of interest to include robustness measures in BO and resolve practical issues such as determining convergence criteria and the initial sample size impact on the optimization results.

CRediT authorship contribution statement

João P.L. Coutinho: Programming, Formal analysis, Conceptualization, Visualization, Writing. **Lino O. Santos:** Methodology, Support to programming, Visualization, Review & editing. **Marco S. Reis:** Conceptualization, Methodology, Review & editing, Resources, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

The authors acknowledge support from the Fundação para a Ciência e Tecnologia (FCT, Portugal), project reference UIDB/00102/2020.

References

Åström, K.J., Hägglund, T., 1984. Automatic tuning of simple regulators with specifications on phase and amplitude margins. Automatica 20 (5), 645–651.

Åström, K.J., Hägglund, T., 2006. Advanced PID Control. ISA, Research Triangle Park, NC.

Åström, K.J., Wittenmark, B., 2011. Computer Controlled Systems: Theory and Design, third ed. Prentice Hall, Upper Saddle River, NJ.

Berner, J., Hägglund, T., Åström, K.J., 2016. Asymmetric relay autotuning – practical features for industrial use. Control Eng. Pract. 54, 231–245. http://dx.doi.org/10.1016/j.conengprac.2016.05.017.

Berner, J., Soltesz, K., Hägglund, T., Åström, K.J., 2018. An experimental comparison of PID autotuners. Control Eng. Pract. 73, 124–133. http://dx.doi.org/10.1016/j.conengprac.2018.01.006.

Brochu, E., Cora, V.M., de Freitas, N., 2010. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv:1012.2599.

Campestrini, L., Eckhard, D., Chía, L.A., Boeira, E., 2016. Unbiased MIMO VRFT with application to process control. J. Process Control 39, 35–49. http://dx.doi.org/10.1016/j.jprocont.2015.12.010.

Campi, M., Lecchini, A., Savaresi, S., 2002. Virtual reference feedback tuning: A direct method for the design of feedback controllers. Automatica 38 (8), 1337–1346. http://dx.doi.org/10.1016/S0005-1098(02)00032-8.

Chen, D., Seborg, D.E., 2002. Multiloop PI/PID controller design based on Gershgorin bands. IEEE Proc. D 149 (1), 68–73. http://dx.doi.org/10.1049/ip-cta:20020092.

Chien, I.-L., Huang, H.-P., Yang, J.-C., 1999. A simple multiloop tuning method for PID controllers with no proportional kick. Ind. Eng. Chem. Res. 38 (4), 1456–1468. http://dx.doi.org/10.1021/ie980595v.

Dittmar, R., 2015. Decentralized SISO active disturbance rejection control of the Newell-Lee forced circulation evaporator. IFAC-PapersOnLine 48 (8), 409–414. http://dx.doi.org/10.1016/j.ifacol.2015.09.002.

Dittmar, R., Gill, S., Singh, H., Darby, M., 2012. Robust optimization-based multi-loop PID controller tuning: A new tool and its industrial application. Control Eng. Pract. 20 (4), 355–370. http://dx.doi.org/10.1016/j.conengprac.2011.10.011.

Euzebio, T.A.M., Silva, M.T.D., Yamashita, A.S., 2021. Decentralized PID controller tuning based on nonlinear optimization to minimize the disturbance effects in coupled loops. In: IEEE Access : Practical Innovations, Open Solutions. Vol. 9. pp. 156857–156867. http://dx.doi.org/10.1109/ACCESS.2021.3127795.

Fiducioso, M., Curi, S., Schumacher, B., Gwerder, M., Krause, A., 2019. Safe contextual Bayesian optimization for sustainable room temperature PID control tuning. In: Proceedings of the Twenty-Eighth International Joint Conference on Artificial Intelligence. International Joint Conferences on Artificial Intelligence Organization, Macao, China, pp. 5850–5856. http://dx.doi.org/10.24963/ijcai.2019/811.

Fröhlich, L.P., Klenske, E.D., Daniel, C.G., Zeilinger, M.N., 2019. Bayesian optimization for policy search in high-dimensional systems via automatic domain selection. In: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems. IROS, IEE, Macau, China, pp. 757–764, http://ieeexplore.ieee.org/abstract/document/8967736.

Greenhill, S., Rana, S., Gupta, S., Vellanki, P., Venkatesh, S., 2020. Bayesian optimization for adaptive experimental design: a review. In: IEEE Access : Practical Innovations, Open Solutions. Vol. 8. pp. 13937–13948. http://dx.doi.org/10.1109/ACCESS.2020.2966228.

Grimholt, C., Skogestad, S., 2018. Optimal PI and PID control of first-order plus delay processes and evaluation of the original and improved SIMC rules. J. Process Control 70, 36–46. http://dx.doi.org/10.1016/j.jprocont.2018.06.011.

Halevi, Y., Palmor, Z., Efrati, T., 1997. Automatic tuning of decentralized PID controllers for MIMO processes. J. Process Control 7 (2), 119–128. http://dx.doi.org/10.1016/S0959-1524(97)82769-2.

Hjalmarsson, H., 2002. Iterative feedback tuning - an overview. Internat. J. Adapt. Control Signal Process. 16 (5), 373–395. http://dx.doi.org/10.1002/acs.714.

Hovd, M., Skogestad, S., 1993. Improved independent design of robust decentralized controllers. J. Process Control 3 (1), 43–51.

Hovd, M., Skogestad, S., 1994. Sequential design of decentralized controllers. Automatica 30 (10), 1601–1607.

Isermann, R., 1989. Digital Control Systems. Springer Berlin Heidelberg, Berlin, Heidelberg, http://dx.doi.org/10.1007/978-3-642-86417-9.

Jones, D.R., Schonlau, M., Welch, W.J., 1998. Efficient global optimization of expensive black-box functions. J. Global Optim. 13, 455–492.

Khosravi, M., Behrunani, V., Myszkorowski, P., Smith, R.S., Rupenyan, A., Lygeros, J., 2022a. Performance-driven cascade controller tuning with Bayesian optimization. IEEE Trans. Ind. Electron. 69 (1), 1032–1042. http://dx.doi.org/10.1109/TIE.2021.3050356.

Khosravi, M., Eichler, A., Schmid, N., Smith, R.S., Heer, P., 2019. Controller tuning by Bayesian optimization an application to a heat pump. In: 2019 18th European Control Conference. ECC, IEEE, Naples, Italy, pp. 1467–1472. http://dx.doi.org/10.23919/ECC.2019.8795801.

Khosravi, M., Koenig, C., Maier, M., Smith, R.S., Lygeros, J., Rupenyan, A., 2022b. Safety-aware cascade controller tuning using constrained Bayesian optimization. IEEE Trans. Ind. Electron. 1. http://dx.doi.org/10.1109/TIE.2022.3158007.

Konig, C., Turchetta, M., Lygeros, J., Rupenyan, A., Krause, A., 2021. Safe and efficient model-free adaptive control via Bayesian optimization. In: 2021 IEEE International Conference on Robotics and Automation. ICRA, IEEE, Xi'an, China, pp. 9782–9788. http://dx.doi.org/10.1109/ICRA48506.2021.9561349.

Lee, P.L., Newell, R.B., Sullivan, G.R., 1989. Generic model control - A case study. Canadian J. Chem. Eng. 67 (3), 478–484. http://dx.doi.org/10.1002/cjce.5450670320.

Liu, T., Wang, Q.-G., Huang, H.-P., 2013. A tutorial review on process identification from step or relay feedback test. J. Process Control 23 (10), 1597–1623. http://dx.doi.org/10.1016/j.jprocont.2013.08.003.

Loh, A.P., Hang, C.C., Quek, C.K., Vasani, V.U., 1993. Autotuning of multiloop proportional-integral controllers using relay feedback. Ind. Eng. Chem. Res. 32 (6), 1102–1107.

Lu, Q., González, L.D., Kumar, R., Zavala, V.M., 2021. Bayesian optimization with reference models: A case study in MPC for HVAC central plants. Comput. Chem. Eng. 154, 107491. http://dx.doi.org/10.1016/j.compchemeng.2021.107491.

Luyben, W.L., 1986. Simple method for tuning SISO controllers in multivariable systems. Ind. Eng. Chem. Process Des. Dev. 25 (3), 654–660.

Makrygiorgos, G., Bonzanini, A.D., Miller, V., Mesbah, A., 2022. Performance-oriented model learning for control via multi-objective Bayesian optimization. Comput. Chem. Eng. 162, 107770. http://dx.doi.org/10.1016/j.compchemeng.2022.107770.

Marco, A., Berkenkamp, F., Hennig, P., Schoellig, A.P., Krause, A., Schaal, S., Trimpe, S., 2017. Virtual vs. real: trading off simulations and physical experiments in reinforcement learning with Bayesian optimization. In: 2017 IEEE International Conference on Robotics and Automation. ICRA, IEEE, Singapore, Singapore, pp. 1557–1563. http://dx.doi.org/10.1109/ICRA.2017.7989186.

Marler, R.T., Arora, J.S., 2010. The weighted sum method for multi-objective optimization: New insights. Struct. Multidisc. Optim. 41 (6), 853–862. http://dx.doi.org/10.1007/s00158-009-0460-7.

McKay, M.D., Beckman, R.J., Conover, W.J., 1979. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. Technometr. J. Stat. Phys. Chem. Eng. Sci. 21 (2), 239–245.

Neumann-Brosig, M., Marco, A., Schwarzmann, D., Trimpe, S., 2020. Data-efficient autotuning with Bayesian optimization: an industrial control study. IEEE Trans. Control Syst. Technol. 28 (3), 730–740. http://dx.doi.org/10.1109/TCST.2018.2886159.

O'Dwyer, A., 2009. Handbook of PI and PID Controller Tuning Rules, third ed. London, Hackensack, NJ.

Ogunnaike, B.A., Ray, W.H., 1994. Process Dynamics, Modeling, and Control. In: Topics in Chemical Engineering, Oxford University Press, New York.

Piga, D., Forgione, M., Formentin, S., Bemporad, A., 2019. Performance-oriented model learning for data-driven MPC design. IEEE Control Syst. Lett. 3 (3), 577–582. http://dx.doi.org/10.1109/LCSYS.2019.2913347.

Rasmussen, C.E., Williams, C.K.I., 2006. Gaussian Processes for Machine Learning. In: Adaptive Computation and Machine Learning, MIT Press, Cambridge.

Rojas, J.D., Flores-Alsina, X., Jeppsson, U., Vilanova, R., 2012. Application of multivariate virtual reference feedback tuning for wastewater treatment plant control. Control Eng. Pract. 20 (5), 499–510. http://dx.doi.org/10.1016/j.conengprac.2012.01.004.

Schillinger, M., Hartmann, B., Skalecki, P., Meister, M., Nguyen-Tuong, D., Nelles, O., 2017. Safe active learning and safe Bayesian optimization for tuning a PI-controller. IFAC-PapersOnLine 50 (1), 5967–5972. http://dx.doi.org/10.1016/j.ifacol.2017.08.1258.

Shahriari, B., Swersky, K., Wang, Z., Adams, R.P., de Freitas, N., 2016. Taking the human out of the loop: a review of Bayesian optimization. Proc. IEEE 104 (1), 148–175. http://dx.doi.org/10.1109/JPROC.2015.2494218.

Shen, S.-H., Yu, C.-C., 1994. Use of relay-feedback test for automatic tuning of multivariable systems. AIChE J. 40 (4), 627–646. http://dx.doi.org/10.1002/aic.690400408.

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

Snoek, J., Larochelle, H., Adams, R.P., 2012. Practical Bayesian optimization of machine learning algorithms. In: Pereira, F., Burges, C., Bottou, L., Weinberger, K. (Eds.), Advances in Neural Information Processing Systems. Vol. 25. Curran Associates, Inc.

Sorourifar, F., Makrygirgos, G., Mesbah, A., Paulson, J.A., 2021. A data-driven automatic tuning method for MPC under uncertainty using constrained Bayesian optimization. IFAC-PapersOnLine 54 (3), 243–250. http://dx.doi.org/10.1016/j.ifacol.2021.08.249.

Sumana, C., Venkateswarlu, C., 2010. Genetically tuned decentralized proportional-integral controllers for composition control of reactive distillation. Ind. Eng. Chem. Res. 49 (3), 1297–1311. http://dx.doi.org/10.1021/ie9008474.

Vu, T.N.L., Lee, M., 2010. Independent design of multi-loop PI/PID controllers for interacting multivariable processes. J. Process Control 20 (8), 922–933. http://dx.doi.org/10.1016/j.jprocont.2010.06.012.

Xue, Y., Li, D., Gao, F., 2010. Multi-objective optimization and selection for the PI control of ALSTOM gasifier problem. Control Eng. Pract. 18 (1), 67–76. http://dx.doi.org/10.1016/j.conengprac.2009.09.004.

Ziegler, J.G., Nichols, N.B., 1942. Optimum settings for automatic controllers. Trans. ASME 64 (759).