Reduced-order controller design via iterative identification and control

Antonio Sala and Alicia Esparza
Dept. of Systems Engineering and Control
Valencia Technical University
Apdo. 22012, E-46071 Valencia, SPAIN
E-mail: {asala, alespei}@isa.upv.es

Abstract

In this paper, a customised version of iterative identification and control algorithms is presented, oriented to the design of reduced-order controllers with increasing performance (in terms of bandwidth for reference tracking). Although a particular set of identification and control design methodologies is chosen due to its simplicity and wide availability, some of the ideas can be applied to alternative methods. The procedure is based on estimating reduced-order models on narrow control-oriented frequency bands when the closed loop performs badly, and interpolating those models and previous ones allowing larger errors at frequencies unimportant for control purposes, so that reduced-order models can still be used. The allowable error for the interpolated model will be determined by commonly used small-gain inequalities.

Keywords: least-squares estimation, control-oriented identification, low-order controllers, robust control, iterative methods.

List of acronyms: high-frequency (HF), identification (ID), low-frequency (LF), least squares (LS), linear time-invariant (LTI), output error (OE), proportional-integral (PI), right half-plane (RHP), robust stability (RS), robust performance (RP), single-input single-output (SISO).

1 Introduction

Experimental design of a controller for an a priori unknown plant implies selecting a “sufficiently high” model order so that it can capture the relevant dynamics up to the desired bandwidth, gather “enough” data to identify the plant with “low” error, and design a “robust enough” controller. However, there are several issues that hinder the success in practice of “one-shot” textbook approaches to ID and control. Those are, for instance, determining the needed model complexity, sample length and optimal excitation signal and determining suitable error bounds for controller design.

Robust control techniques [24] can yield satisfactory results if combined with suitable model order selection, ID experiment design [15, 6, 4] and error bound estimation techniques ([9, 18, 5]).

However, many techniques depend on characteristics of the plant that are unknown at the beginning, so after some information is gathered the criteria are changed. Furthermore, when a controller is designed, it reduces sensitivity to modelling errors at some frequencies but enhances it at others so it asks for a recasting of the ID criteria. The use of the information gathered in unsatisfactory closed-loop experiments gives rise to iterative ID and control approaches [7, 19, 22, 1], where the controller tuning operation is a series of interleaved ID experiments and controller improvement steps (limited in size by estimates of stability margins).

An additional issue arising in many situations is the need of reduced-complexity controllers. The use of complex models to design high-performance regulators meets with many difficulties, due to mainly these
reasons: numerical precision (high-order polynomials need very precise coefficients [16, 11]), noise (high-bandwidth regulators have a big HF gain, producing unacceptable actuator activity even if model were correct; poor signal-to-noise ratio produces heavy parameter variance for high-order models) and nonlinearity (Simple linear models might suffice for increasing performance until actuator and rate saturation, plus other nonlinearities hinder performance improvement).

There are different alternatives to design of reduced-complexity controllers. In this work, a reduced-order-model approach is sought. As high-order models cannot be successfully identified in many practical situations and, even if they could, as feedback reduces sensitivity to modelling errors, lower-order ones could suffice: the catch lies in concentrating the model fit at frequencies critical to control robustness. An iterative methodology to design reduced-order controllers for an a priori unknown plant is presented, based on estimation of reduced-order models on particular control-relevant frequency bands.

The use of an iterative approach arises from the fact that the control-relevant frequencies are not known a priori with enough accuracy, as they depend on both the control goals and the plant. The performance goals will be progressively increased, as in the windsurfer methodology [2], starting from low-bandwidth, low-authority controllers. In this work, different model parameterisations and order-increase criteria are used, and performance assessment is based on quality of step response. Reduced order models are adjusted to fit both the last experimental data batch and previous lower-frequency models, combining loss-function analysis and small-gain inequalities. If no suitable fit can be found, the model order is increased. Using LF information from previous designs allows to determine ample bounds for the LF modelling error in future models. As the controller must be robust enough to cope with the HF uncertainty, the procedure works reasonably well if the plant has a HF roll-off, otherwise it is the controller who should provide that roll-off, increasing its order1.

The structure of the paper is as follows: first, section 2 presents basic definitions regarding ID, robust control, and iterative ID and control. Section 3 discusses iterative design of reduced-order controllers. Section 4 details examples of the methodology. Finally, some conclusions summarise the main ideas.

2 Preliminaries. Basic Definitions

On the following, it will be assumed that a controller must be designed for a discrete-time SISO LTI process described by the equation:

\[ y(t) = G_s(q)u(t) + v(t) \]  \hspace{1cm} (1)

where \( q \) stands for the delay shift operator and \( v(t) \) is an independent coloured-noise disturbance sequence.

A parameterised model for the process will be denoted as:

\[ y_m(t) = G_\theta(q)u(t) + v_m(t) \]
\[ v_m(t) = H_\theta(q)e(t) \]  \hspace{1cm} (2)

where \( e(t) \) is a white noise and \( H_\theta \) is the disturbance model. Denoting as \( r(t) \) the desired system trajectory, the closed-loop controller configuration will be the usual one:

\[ u(t) = K(q)(r(t) - y(t)) \]

so the nominal transfer functions from disturbance to output (sensitivity \( S \)) and from setpoint reference to output (complementary sensitivity \( T \)) are given by \( S = (1 + G_\theta K)^{-1}, T = 1 - S \). On the following \( G \) will be used as shorthand for the parameterised plant model \( G_\theta \). The achieved sensitivity and complementary sensitivity will be denoted as \( S_a, T_s \), respectively.

1The benchmark plant has a double differentiator plus resonance, i.e., a significant upward slope in frequency domain. Furthermore, usual step response measures lose part of its interpretability in a differentiating plant. Due to these reasons, the proposed procedure could not find a satisfactory low-order controller for the benchmark plant.
2.1 Uncertainty and feedback

Model-based design provides a candidate controller based on a nominal model \( G_\theta \). A good feedback design can achieve satisfactory performance even on the face of significant differences between the nominal and real plants, even accommodating nonlinearities. In fact, the uncertainty-reducing effect of feedback can relax the requirements for ID. The following results determine some bounds of unstructured modelling errors for RS and RP.

**Lemma 1 ([24])** Let \( \Pi \) be the set of plants given by \( \Pi = G + \Delta \), \( \Delta \in \mathcal{RH}_\infty \) (stable proper rational transfer functions), and let \( K \) be a stabilizing controller for the nominal plant \( G \). Then, the closed-loop system is well-posed and internally stable for all plants in \( \Pi \) verifying \( |\Delta(j\omega)| < |K(j\omega)S(j\omega)|^{-1} \), where \( S = (I + GK)^{-1} \).

So, some frequencies allow more uncertainty than others without compromising closed-loop stability, i.e., in zones where \( KS \) is small the uncertainty size can be bigger. Other uncertainty descriptions (multiplicative, coprime factor, etc.) are widely used and similar formulae can be obtained for them [20]. The modulus is replaced with maximum singular value norm for transfer function matrices.

Regarding RP, it can be shown that:

**Lemma 2 ([24])** For a system with additive uncertainty \( \Delta \), \( |\Delta| < |w_u| \) the achieved sensitivity \( S_\epsilon \) will be below a desired RP bound \( w_p^{-1} \) if the nominal sensitivity verifies: \( |w_pS| + |KSw_u| < 1 \)

**Lemma 3** If additional phase information of \( \Delta \) is available, if nominal performance and stability hold and \( -\pi/2 \leq \arg(KS\Delta) \leq \pi/2 \) for every frequency in which \( |w_pS| + |KS\Delta| > 1 \) then the closed loop is stable and RP does also hold.

**Proof:** If uncertainty is \( |w_pS| + |KS\Delta| < 1 \) then Lemma 2 applies. Else, \( S_\epsilon^{-1} = |1 + K(G + \Delta)| = |1 + KG + K\Delta| = |1 + KG||1 + KS\Delta| \geq |1 + KG| = S^{-1} \) as the factor \( 1 + KS\Delta \) has real part greater or equal to 1 from the conditions of the lemma (real part of \( KS\Delta \geq 0 \)). As nominal performance holds, \( S < w_p^{-1} \), so \( S_\epsilon < S < w_p^{-1} \) and RP holds. RS follows from straightforward considerations on the Nyquist diagram (to encircle the point \(-1, KS\Delta \) must be allowed to take negative real values).

2.2 Identification for control

Given a data set \([y(t),u(t)]\), the prediction error [15] associated to model (2) is:

\[
p(t) = H_\theta^{-1}(q) \left( y(t) - G_\theta(q)u(t) \right)
\]  

(3)

Least squares (LS) algorithms try to find a parameter vector \( \theta \) that minimises \( V(\theta) = \sum_{t=1}^{N} e(t)^2 \) where \( e(t) \) is a sequence of N prediction errors generated by filtering \( p(t) \) by a user-defined prefilter \( L(q) \). The direct closed-loop LS algorithms try to estimate a model of the plant from plant input and output data sets obtained with a controller in operation.

**Lemma 4 ([15])** For long data lengths \((N \to \infty)\), the minimisation index used by direct closed-loop algorithms, in frequency-domain, asymptotically approaches:

\[
V(\theta) = \int_{-\pi}^{\pi} \left[ \left| G(e^{j\omega}) - G_\theta(e^{j\omega}) \right|^2 |\Phi_r(\omega) + \frac{1}{K(e^{j\omega})} + G_\theta(e^{j\omega})|^2 |\Phi_v(\omega) \right] \frac{|K(e^{j\omega})|^2 |L(e^{j\omega})|^2}{|1 + G(e^{j\omega})K(e^{j\omega})|^2 |H_\theta(e^{j\omega})|^2} d\omega.
\]

(4)

where \( \Phi_r(\omega) \) and \( \Phi_v(\omega) \) are the frequency spectrum of the reference \( r \) and the noise \( v \) respectively. The noise is assumed not measurable.

There are alternative closed-loop ID methodologies that try to remove the noise-induced bias term [14, 15, 21], to be applied in processes with significant disturbance effect. The estimated parameters also have noise-induced variance. Under restrictive assumptions (plant should be in the model set) variance is asymptotically proportional to the number of parameters and inverse proportional to the data length and signal-to-noise ratio. Closed-loop expressions indicate that variance increases with regulator “gain” [15, 8].
Iterative identification and control. In a model-based control design, the designer calculates a nominal closed-loop system and then tries the resulting controller on the actual plant, so that a discrepancy occurs. As a result, an iterative approach originates: get a model, design a controller and then identify a new model from the closed-loop test. The question is if those iterations converge to a controller achieving a particular user-defined sensitivity. The question was analyzed in [10] and the answer was, in a general case, negative, as (a) LS minimisation relates to the 2 norm (integral of the frequency response) and not the ∞-norm (peak of frequency response) used in lemma 1; (b) the controller and sensitivity functions themselves are function of θ in model-based design: changing the controller can invalidate a previously good model.

However, iterative ID and control has had a significant impact [1] as, in practice, a controller designed based on a model will not usually achieve “optimal” performance on a real plant in its first try: when put into operation, closed-loop data will be available to refine its behaviour, by improving model quality as more data are gathered. This batch-like processing has, obviously, a relationship with adaptive control [13] where parameter update is done after every sample.

To ensure a final acceptable controller, caution measures must be put in place: correct ID experiment design (excitation, data selection, prefilter, number of model parameters, model validation) and appropriate “step size” control in the iteration process (limiting the difference between successive models or controllers). For further detail, the reader is referred to [1, 3, 7, 22].

The windsurfer approach [2], for example, is based on starting with low-bandwidth low-authority controllers, and then progressively increasing specifications until unacceptable discrepancies with the designed behaviour appear. At that point, re-identification minimising a control-oriented criterion is carried out so model is refined to allow design of higher performance controllers. On the sequel, based on these ideas, a scheme using widely available algorithms will be proposed, with the goal of keeping model order low as long as controllers designed based on it do exhibit acceptable performance.

Experiment Design. As it is clear now, use of LS algorithms in ID for control in practice needs a careful experiment design to obtain meaningful results [15, 4]:

Parsimony. Complex (high-order) models entail complex controllers. Furthermore, to diminish variance, models with few parameters should be used, but that restricts average model fit capabilities (bias). If model fit is concentrated on robustness-critical frequencies, large errors can be tolerated at other frequency ranges.

Disturbance models. Frequently, in ID for control one can choose high-order disturbance models to sweep away correlations [17] as errors in its parameters will not destabilise the closed-loop system.

Filtering and excitation. The prefilter L in (4) should under-weight regions of the frequency spectrum dominated by disturbances and enhance control-oriented ones (excited by a suitably chosen input signal).

3 Reduced-order Iterative Algorithm

In this section, previous ideas on uncertainty bounds, structured uncertainty and experiment design in ID for control will be applied to the design of low-order regulators achieving high-performance control (in terms of reference tracking). The procedure will apply to stable SISO linear plants.

Algorithm 1 (Reduced order iterative ID and control)
The summary of the proposed procedure is (see later for details):

1. Start from a conservative, very-low-bandwidth controller for a model given by a reasonable estimate of the plant’s LF behaviour (for example, PI controller from a DC gain estimate).

2. Performance evaluation: increase performance specifications (closed-loop bandwidth) until the observed performance starts to decline.
3. Identify a low-order model fitting a control-relevant frequency band around the frequencies where performance was unacceptable.

4. Design a controller in the bandwidth range that previously failed, with a conservative HF roll-off to account for the not-yet-excited HF uncertainty.

5. Model-controller validation: to design a controller with a particular bandwidth some bounds on the modelling error have to be met (previous models and its validity regions are used to evaluate LF fit). If validation is unsuccessful, re-identify with a changed criterion. If no valid model can be found then increase model order.

6. Implement the previously designed controller and go either to step 2 if the real loop behaves successfully or to step 3 if it does not.

**Termination.** The algorithm will end due to one of three reasons: when satisfactory performance is reached, when significant actuator saturation is present (unless the control design methodology takes that into account) or when several steps of model-order increase do not yield an improvement on the measured performance.

In practice, the algorithm would need to be stopped if nonlinearity and low signal-to-noise ratio render ID at high frequencies impossible.

There are various choices of ID, model validation, controller design methodologies and performance evaluation, so that the previous scheme actually refers to a family of algorithms. A particular choice of techniques will be made but it is in the authors’ opinion that a significant fraction of the conclusions and results could be extended to alternative techniques.

### 3.1 Design methodologies

The methodologies proposed for the relevant steps of the algorithm will now be detailed, applied to sampled-data SISO linear systems.

**Evaluation of the achieved performance (Step 2).** On the real plant, the decision on whether the achieved performance of the current closed-loop is satisfactory can be based on many criteria (quality of step response, estimation of achieved peak of a particular closed-loop function \( S, T, KS, \ldots \)). In our case, mostly motivated by the simplicity in the automation of the procedure, a model \( G_j \) will be experimentally invalidated when the closed-loop with the controller designed using \( G_j \) exhibits an overshoot in step response greater than a pre-specified bound or noticeable HF oscillations occur in input or output signals.

Note that there are two different types of “bad” real performance:

- **HF misfit** (unacceptable overshoot and oscillations in signals, pointing out that stability limits are being approached). There is a need to re-identify the plant around the frequencies determined by those oscillations. They are pinpointed when its severity increases when (cautiously) increasing bandwidth.

- **LF misfit** (slow correction of position errors, for example). The solution to that has two alternatives: either having a better LF model or increasing the gain of the regulator (by increasing the designed bandwidth if there are no significant stability problems). In this work, the second choice will be preferred as it will allow simpler models with large errors in the LF range. Obviously, the increase in bandwidth is simultaneous to an increase in measurement noise amplification (approximately at the same rate than the inverse of the plant) so, in real plants, reaching impractical limits in these amplifications would ask for a change in the control design philosophy, trying to estimate better LF models. Obviously, that will likely lead to higher-order models.
Identification (Step 3). As previously pointed out, the chosen ID method will be a LS estimation [15] of an OE model \( y = \frac{B}{A}u + v \) from input-output data gathered in closed-loop operation so (4) applies. Reference excitation is injected into the loop generating a multi-sine signal (random phase shift) with frequencies linearly spaced on the required frequency band (see later).

OE algorithm implementations yield always a stable system so the minimisation in (4) is only made over the set of stable LTI plants of the desired order. As the issue is not estimation of the “true” plant but an approximation of a zone in the frequency domain, better approximations could be achieved if the search was made also on the unstable parameter space, i.e., the counter-intuitive possibility that an unstable reduced-order model for a stable plant works best cannot be discarded. However, it will not be pursued any further, as conditions in lemma 1 do not allow for a different number of RHP poles in plant and models, so the uncertainty representation would need to be changed.

At this point, there are several important choices to be made:

a) Identification frequency bands.

ID will be performed to fit those frequencies where performance is poor due to stability limits being approached. As mentioned above, these are the ones which produce overshoot and manifest oscillations in input or output signals. The frequency of those oscillations can be determined from the step response. Once it is known, additional excitation on a band around it is injected for ID purposes. The maximum and minimum frequencies can be defined by determining the oscillation frequency \( \omega_{id} \) and defining a factor \( \gamma \) so the ID frequency band is \([\omega_{id}/\gamma; \omega_{id} \times \gamma]\). As an alternative, if it were possible to carry out closed-loop frequency response tests, the band could range from the frequency of the oscillations (approximately the peak of \( T_c \) or \( KS_c \)) to the frequency in which the magnitude of the corresponding closed-loop function had decreased by a given amount (say, for example, 10 dB).

If the chosen band is too narrow (consider nearly a point in the frequency domain), the updated model will not be useful for extrapolation to higher frequencies, i.e., for achieving a significant bandwidth increase. On the other hand, if the band is too wide, extra parameters might be needed to identify features that might be irrelevant for control purposes at this moment and noise-induced estimated parameter variance will be higher. So, there is a tradeoff between sticking to the lowest-possible model order (with a likely increase in the number of iterations) or identifying a higher-order model able to withstand substantial bandwidth increases (say, more than half a decade).

b) Model order.

Once data from a particular frequency band is gathered, the choice of model order will be made so that the prediction error variance is small enough when compared to overall output variability, or in the sense that an increase of the number of parameters does not produce a significant reduction of that variance [15]. As previously mentioned, the wider the excited frequency band is, the bigger the needed order will be under these criteria. Note that non-minimum-phase zeroes usually appear in the reduced-order models as artifacts to achieve a better fit in the frequency response band under study.

Controller design (Step 4). Various controller design methodologies can be used. In the examples in section 4, a stacked sensitivity approach [20, 24] is used, where a minimisation of the \( \infty \)-norm of:

\[
N = \begin{bmatrix} w_p S & w_T T & w_U KS \end{bmatrix}^T
\]

is done over the set of internally stabilizing controllers. \( w_p \) is a low-pass weight penalising performance, \( w_T \) penalises the complementary sensitivity (achieving robustness to relative modelling error) and \( w_U \) penalises the disturbance-to-control transfer function (robustness to additive error). The last two weights are usually designed to have significant impact at high frequencies. To get reduced order controllers, reduced order weights should be used, such as constant or first order.
In initial steps of the procedure, low-performance conservative controllers are designed. As bandwidth increases, the robustness weights need to be decreased for a solution to exist (performance-robustness tradeoff) so smaller modelling error bounds are tolerated. As iterations progress, the greater precision needed is usually concentrated at higher frequencies, and model mismatch will actually yield unsatisfactory performance pointing out the need of re-identification. Note that the extent of the needed weight modifications is an indicator of how cautious the bandwidth-increase step is.

Another possibility is to directly specify the complementary sensitivity $T$ as:

$$T = G_{nmp} \frac{(1 - e^{-T \lambda})^n}{(z - e^{-T \lambda})^n}$$

with increasing nominal bandwidth $\lambda$, being $n$ at least the model relative degree. Negative-real model zeros and those outside the unit circle are placed on the numerator of $T$ in the factor $G_{nmp}$, to avoid undesirable pole-zero cancellations. The implemented controller in that case is $T G_{\theta}$, where $G_{\theta}$ is the current reduced-order plant model.

**Model-controller validation (Step 5).** Let us consider a new model $G_j$, identified in the frequency band $[w_j^-, w_j^+]$. Although that frequency band will be quite likely the most control-relevant zone of the frequency spectrum, the validity of that model has to be checked against the past data records available at the moment: by designing controllers with increasing bandwidth, the historic record of past experiments has a lot of information about the LF behaviour of the plant. In fact, there is a record of data and a list of models obtained from them $G_0, \ldots, G_{j-1}$, each of them meant to fit a particular frequency band $B_k = [\omega_k^-, \omega_k^+]$, with $0 \leq k \leq j - 1$. In a certain sense, then, as the real plant will be near each of the models, it is a kind of a structured description of the LF modelling error.

Let us assume that $G_k$ is a reasonable fit to the real plant between frequencies $[\omega_k^-, \omega_k^+]$, so that quality of a new model $G_j$ will be tested by comparing it to $G_k$. An error bound for the fit of $G_k$ will be given by the following lemma.

**Lemma 5** For each frequency $\omega$, let $G_k$ be the best approximation of the plant $G_s$ at it (so $\omega \in [\omega_k^-, \omega_k^+]$). Then, if a model $G_j$ verifies:

$$|G_s - G_k| < \left| \frac{1 + G_j C}{C} \right| - |G_k - G_j|$$

conditions of Lemma 1 are verified at $\omega$.

**Proof:** For RS, designing a controller $C$ with model $G_j$, the following inequality should hold at all frequencies:

$$|G_s - G_j| \frac{C}{1 + G_j C} < 1$$

but, for any $k$:

$$|G_s - G_j| = |(G_s - G_k) + (G_k - G_j)|$$

and all $G_s, G_k$ and $G_j$ are stable (due to the OE algorithm). For a particular frequency $\omega$, applying the triangle inequality:

$$|G_s - G_j| \frac{C}{1 + G_j C} \leq (|G_s - G_k| + |G_k - G_j|) \frac{C}{1 + G_j C}$$

So, if $G_s, G_k$ and $G_j$ verify

$$|(G_s - G_k| + |G_k - G_j|) \frac{C}{1 + G_j C} < 1$$

solving for $|G_s - G_k|$ expression (7) is obtained.
The obvious choice is to try to identify a $G_k$ that gives the smallest $|G_s - G_k|$. The following corollary expresses a conservative bound for the required ID accuracy of $G_k$ as a function of how current $G_j$ fits $G_k$.

**Corollary 1** If, in the validity range of past model $G_k$, the following inequality is satisfied for current $G_j$

$$|G_k - G_j| < \alpha_k$$

then, to satisfy the RS condition in lemma 1, in the frequency band $[\omega_k, \omega_{k+}]$, the allowed modelling error for $G_k$ is:

$$|G_s - G_k| < (1 - \alpha_k) \frac{1 + G_j C}{C}$$

The corollary implies that the bigger the allowed discrepancy between the current reduced-order model $G_j$ and the best approximation $G_k$ to the true plant at a particular frequency, then the more precise the approximation $G_k$ needs to be. However, forcing $\alpha_k$ to be small for all $k$ would require a greater complexity (higher order) for $G_j$.

The previous results give a bound for modelling error that motivates the following model validation test.

**Algorithm 2 (Validation test.)**

For all past models $G_k$, if the last estimated model $G_j$ and a controller designed on it verify:

$$\max_{\omega \in [\omega_k, \omega_{k+}]} |(G_k - G_j) \frac{C}{1 + G_j C}| < \alpha_k,$$

at the frequencies $[\omega_k, \omega_{k+}]$ where $G_k$ was identified, being $0 < \alpha_k < 1$, the model $G_j$ is deemed valid and the controller is tested on the real plant. The maximum in (11) can be obtained by calculation of the $\mathcal{H}_\infty$ norm of the transfer function multiplied by a band-pass filter. $\alpha_k$ are user-defined parameters related to the quality of previous models. As lower signal-to-noise ratio is present at high frequencies, expected quality of the models decreases. So the higher the ID frequencies are the smaller $\alpha_k$ is recommended.

Based on Lemma 3, if (11) is not verified for a particular $G_k$ but the argument of $(G_k - G_j) \frac{C}{1 + G_j C}$ is in the range $[-\pi/2, \pi/2]$, the model-controller pair can be validated as well. That will be used to partly reduce conservativeness at LF by tolerating larger errors.

Regarding the HF bands, as they have not been excited, the HF modelling misfit will eventually make the to-be-tested controllers behave poorly, so the need for new ID experiments will become manifest, as intuitively expected.

Obviously, the information of past experiments does not cover all frequency bands. As scarce knowledge is available on some frequencies, in the first test on the true plant of a new controller, a conservative regulator should be tried (cautious step-size in specifications with respect to the last previously working regulator), with a reasonable HF roll-off slope. Obviously, if the plant itself exhibits a reasonable roll-off slope (and it is known a priori) the controller order can be lower. Thus, the HF roll-off of the complementary sensitivity $T$, specified by the design of weight $w_T$, is one of the tuning knobs regarding caution in the presented procedure with respect to the not-yet-manifested HF process dynamics. Another possibility is modifying $w_U$. Note that introducing more caution in weights might lead to higher-order regulators. To try to avoid that situation, the to-be-tested controller is designed for a bandwidth around the left extreme of the frequency band used to identify the new model.
3.1.1 Off-line experiment design in case of invalidation

After the ID procedure the LF modelling error of the new model $G_j$ might be too high so that it is deemed invalid by algorithm 2. That situation may be originated by one of these two conditions: either there is enough model flexibility (number of parameters) but the data contained too little LF excitation, or the number of parameters is not enough to both approximate the new frequency band and the LF ones to a satisfactory degree.

To check which condition applies, a virtual set of data will be synthesised so that it has the necessary excitation profile, by adding to the last data obtained the simulated outputs from previous models subject to multisine input excitation (directly adding past data would corrupt data due to presence of broadband noise).

For each model a data set $Z_k = (y_k, u_k)$, $k = 1, \ldots, j-1$ is generated, where $y_k$ is the simulated plant output (using $G_k$) with $u_k$ being a multisine input signal containing frequencies in the ID band of the model $[\omega_k^-, \omega_k^+]$.

Let us assume that a model $G_j$ (identified in the frequency band $[\omega_j^-, \omega_j^+]$) verifies, for each $k$: \[
\max_{\omega \in [\omega_k^-, \omega_k^+]} |(G_k - G_j) \cdot \frac{C}{1 + CG_j} | = \beta_k, \ 1 \leq k \leq j - 1
\] (12)
so that some of the $\beta_k$ are greater than the user-defined bounds $\alpha_k$ and the argument of $(G_k - G_j) \cdot \frac{C}{1 + CG_j}$ is outside $[-\pi/2, \pi/2]$. Let us denote as $\mathcal{K}$ the set of $k$’s from models in that situation.

In that case, a sequence of “virtual” data $Z^{[m]}$ will be designed to look for a better LF fit. Starting with the data for the last experiment $Z^{[1]} = Z_j$, then:

\[
Z^{[m+1]} = Z^{[m]} + \sum_{k \in \mathcal{K}} \eta(\beta_k - \alpha_k)Z_k
\] (13)
where $\eta$ is a sufficiently small user-defined step size. In this way, data from an invalidating older model (with excitation at the frequencies used to identify $G_k$) is added, with $\eta(\beta_k - \alpha_k)$ as its weighting factor, to the one from the last experiment.

With that new synthetic data, a new model $G'_j$ is obtained, a controller is designed for the target bandwidth, and (12) is calculated again. To check that fitting lower frequencies has not distorted too much the fit at the frequency band of $G_j$, an additional check is made:

\[
\max_{\omega \in [\omega_k^-, \omega_k^+]} |(G_k - G'_j) \cdot \frac{C}{1 + CG'_j} | = \beta_k, \ 1 \leq k \leq j
\] (14)
so that:

- if all $\beta_k$ are below the user-defined bounds $\alpha_k < 1$, or the argument condition is verified, the model $G'_j$ is validated and the controller is tested on the real plant.
- if $\beta_j < \alpha_j$ and some $\beta_k$ are above $\alpha_k < 1$ ($1 \leq k \leq j - 1$), and the argument condition is not verified, a new set of data is generated with (13) so that a new $G'_j$ is identified and tested for validity.
- if $\beta_j \geq \alpha_j < 1$ then improvements in LF behaviour imply unacceptable misfit at the frequencies that are important for control at current bandwidth. In this case, the order of the model must be increased in order to both fit the frequencies of interest and to be validated at lower frequencies. So, the procedure is restarted with a $G'$ of higher order and $Z^{[1]} = Z_j$.

4 Examples

In this section a number of examples are presented to demonstrate the applicability of the proposed procedures.
Example 1: 5th order plant

Given a plant with the following transfer function:

\[ G(s) = \frac{41.67(s + 1)}{(s + 70)(s + 1)(s + 15)(s + 1)(s + 1)(s + 4)(s + 4)(s + 200)(s + 800)(s + 1500) + 1} \]

It is to be controlled using a sampling rate of 1000 Hz, so its discrete-time model is:

\[ G(z) = \frac{0.16686(z + 2.059)(z - 0.9851)(z - 0.9324)(z + 0.1351)(z - 0.996)(z - 0.956)(z - 0.8187)(z - 0.4493)(z - 0.2019)}{(z - 0.9851)(z - 0.9324)(z + 0.1351)(z - 0.996)(z - 0.956)(z - 0.8187)(z - 0.4493)(z - 0.2019)} \]

A white noise of amplitude 0.01 is added to simulate sensor noise.

The controllers are designed via the stacked sensitivity approach (5) with weights \( w_T \) and \( w_U \) being constants. \( w_P \) is a first order transfer function with the following structure:

\[ w_P^{-1} = 0.008 \frac{s}{\lambda^2} + 1 \]

aiming at small sensitivities (-42 dB) at LF bands and a maximum peak of 9 dB. \( w_P \) is discretised via a bilinear transformation. The point where the designed \( S \) gets above 0 dB is approximately 0.4\( \lambda \), the nominal bandwidth (\( S < 6 \) dB) is 0.2\( \lambda \).

The starting model is just the DC-gain of the process \( G_1 = 41.67 \). This model allows to design PI controllers with increasing \( \lambda \) with less than 10% overshoot up to 9 rad/s (see figure 1).

Further increases of \( \lambda \) causes oscillations to appear at a frequency around 4.3 rad/s, so it is decided to identify in closed-loop a new model in the band [2.8, 7.6] rad/s with multisine reference excitation. The loss function (prediction error variance) for a first order model is 0.00294, very small when compared to the output variance \( \sigma_y = 37.0235 \). So a first-order model \( G_2 = \frac{7.9185(z - 0.978)}{z - 0.9958} \) is deemed precise enough. It allows reaching \( \lambda = 170 \) rad/s (see figure 2) in subsequent steps of increasingly stringent controllers.

Figure 1: Iterations with \( G_1 = 41.67 \) up to \( \lambda = 9 \) rad/s.

Figure 2: Iterations with a first order model up to \( \lambda = 170 \) rad/s.
The controller order is greater than that of the model due to the weights in the stacked sensitivity procedures. 

depicted in figure 4 with \( \lambda \) is obtained adding synthetic data to fulfill the validation tests. With this model the plant achieves the response \( T \), as in equation (6). Following a similar methodology as in previous examples, the model \( G_3 \) is preliminarily identified in the frequency range \([340, 560] \) rad/s. As it does not pass the RS tests with \( G_1 \) and \( G_2 \), it is necessary to obtain a new second order model taking into account some LF information, using algorithm 2. The new \( G_4 = \frac{2.2419(z^2 - 2.227z + 2.083)}{(z - 0.9501)(z + 0.5756)} \) is validated successfully against \( G_1 \), \( G_2 \) and \( G_3 \) so that \( \lambda \) could be increased up to 875 rad/s (see figure 3). But this time iterations stop because the control action presents HF oscillations (around 2020 rad/s). Than means that the achieved \( KS \) has a large value around this frequency.

Next model \( G_5 \) is a third order one, as the loss function with this model diminishes considerably. The frequency band for its ID is \([1900, 2100] \) rad/s, to obtain a reasonable estimate of the problematic frequency range. This model cannot be validated against all LF models, so a new \( G_5 = \frac{0.007862(z + 20.27)(z + 2.241)(z - 0.02211)}{(z - 0.9527)(z^2 - 0.7157z + 0.2173)} \) is obtained adding synthetic data to fulfill the validation tests. With this model the plant achieves the response depicted in figure 4 with \( \lambda = 1050 \) rad/s. The final controller is

\[
C = \frac{1.0182(z + 0.8374)(z - 0.9527)(z^2 - 0.7157z + 0.2173)}{(z - 0.9971)(z - 0.02755)(z^2 + 0.797z + 0.3635)}
\]

The controller order is greater than that of the model due to the weights in the stacked sensitivity procedures.

As \( \lambda \) increases further and approaches the Nyquist frequency (3140 rad/s), only marginal improvement of performance is achieved. Further performance improvement would need a faster sampling rate (achieved settling time was approximately 4 samples), increasing model order for better LF accuracy, or changing the designed sensitivity weight increasing its slope and/or its LF attenuation. The additive modelling error of each of the models is depicted in figure 5.

**Alternative control strategy.** The same plant can be controlled by specifying the complementary sensitivity \( T \), as in equation (6). Following a similar methodology as in previous examples, the model \( G_1 = 41.67 \) allows reaching \( \lambda = 4.5 \) rad/s with less than 10% overshoot; \( G_2 = \frac{7.8998(z - 0.9779)}{z - 0.9958} \) reached \( \lambda = 65 \) rad/s; \( G_3 = \frac{-1.2948(z - 1.833)(z - 0.9689)}{(z - 0.9894)(z - 0.8725)} \) reached \( \lambda = 860 \) rad/s and with the 2nd-order model \( G_4 = \frac{0.039027(z^2 + 4.376z + 11.76)}{(z - 0.9513)(z - 0.4531)} \).
the bandwidth could be increased up to the point of placing the designed poles at the origin in the $z$-plane, with the controller:

$$C = \frac{1.588(z - 0.9513)(z - 0.4531)}{(z - 1)(z + 0.7288)}$$

Figure 6 presents the final step response in output and control action.

**Comparison to full-order ID + model reduction.** In an ideal LTI plant case, gathering “enough” data from the system subject to “broadband” excitation can successfully identify a full-order model. In this case, estimation of a 5-th order model with $10^4$ data samples from white-noise input yielded a nearly perfect model (achieving that precision in a realistic situation is harder). With the full-order model, alternative model reduction techniques can be used.

In fact, the proposed methodology can be modified to act as a frequency-weighted model reduction technique: from a full-order model and a candidate controller, a reduced-order model is obtained by identifying a reduced-order model at the most critical frequency band for robustness to additive error (order to be decided by loss-function analysis in a first step). Then, the validation test of algorithm 2 and the subsequent experiment design in case of invalidation can be applied selecting a set of frequency points where the RS and RP inequalities are not fulfilled.

Applying the modified technique to the identified full-order model of the plant under study yielded a 3rd-order model achieving the step response in figure 7. Balanced truncation to 3rd-order also gave similar results.

**Example 2: experimental setup**

The proposed iterative ID and stacked sensitivity control algorithm was applied to a real system consisting of a dc-motor and an aluminium rod joined by springs (Quanser\textsuperscript{TM} rotary flexible joint experiment, modules SRV2 and ROTFLEX\textsuperscript{2}). A first-principle model would have fourth order dynamics plus higher-frequency
modes from electrical time constants and rod flexibility. Nonlinearity comes mainly from saturation, friction in the motor and spring geometry. The actuator was a power amplifier yielding voltages in the range ±10 V. Controllers were implemented in C language. Sampling period was 1 ms. Each experiment gathered 30 s of input-output data for analysis. The first ID experiment was a steady-state speed measurement with two different (positive) voltages (1 V, 2.45 rad/s) and (2 V, 5.39 rad/s). Linear extrapolation to zero obtained a dead-zone of 0.16 V. To achieve better linearity and steady-state error, compensation of that dead-zone was carried out by adding or subtracting its value depending on the sign of the control signal.

**Full-order ID approach.** In [12] a solution is given for the control of the system based on ID of a 4th order model (termed subsequently ”full-order” model) for a sampling period of 30 ms. A polynomial controller $K = R/S$ with order $n_R = 5$ and $n_S = 3$ (9 parameters) has been designed by pole placement with sensitivity function shaping. It achieves on the real system a time response for a step set-point change (from 0 to 0.9 V) of approximately 0.8 s.

Subsequently we will illustrate for the same problem how our iterative approach allows to generate controllers of improved performances starting from a very poor initial model, up to a performance comparable to the full-order approach.

**Iterative approach.** The approach proposed in this paper was applied to the plant under consideration. To design very-low-bandwidth controllers, the slope of the speed-vs.-voltage line in the dead-zone experiment determined a first “gain + integrator” model:

$$G_1 = \frac{2.94 \cdot 10^3}{z - 1}$$

With that model, regulators could be designed up to values of $\lambda = 10$ rad/s. At that bandwidth, the overshoot was deemed excessive and a new (first-order plus integrator) model $G_2$ was estimated on the frequency band $[5,30]$ rad/s (with band-pass filtered white-noise excitation). After validation, the models and controllers are:

$$G_2 = \frac{7.576 \cdot 10^{-4}z - 7.402 \cdot 10^{-4}}{(z - 1)(z - 0.994697)} \quad C = \frac{0.03883z^2 + 0.008559z - 0.046933}{z^2 - 1.9702z + 0.97047}$$

The step response with the first and second models is depicted in figure 8.

The model $G_2$ can be squeezed up to a value of $\lambda = 120$ with satisfactory performance. For $\lambda = 180$ manifest oscillations appear, signaling that stability limits are approached. Responses are shown in figure 9.

A new model $G_3$ is identified in the band $[100,500]$ rad/s. After validation, it still is a 1st-order plus integrator. The second-order controller designed on it is

$$C_3 = \frac{14.84z^2 + 2.585z - 12.33}{z^2 - 0.9567z + 0.2128}$$

achieves a final response with $\lambda = 240$ depicted in figure 11. Note that marginal improvement is achieved over the previous design and measurement noise amplification is bigger. Control actions for $< G_2, \lambda = 120 >$ and $< G_3, \lambda = 240 >$ are depicted in figures 10 and 12.
Figure 8: $\lambda = 10$. Changing from $G_1$ to $G_2$. Response to a 2 V step at $t = 3$.

Figure 9: Behaviour with the second model.

The best performances achieved with $G_1$, $G_2$ and $G_3$ are depicted in figure 13. Settling time is about 0.6-0.7 s. Further ID experiments in different bands couldn’t improve achieved performance anymore.

Bandwidth increases in real systems are not as easy to achieve as in simulations: in the authors’ opinion, performance is limited by nonlinearities and signal-to-noise ratio. In fact, saturation is present in a significant fraction of the transient and regulators trying to squeeze performance while having significant LF errors result in increasingly higher measurement noise amplification. Furthermore, sinusoidal input experiments showed that at 20 rad/s with input amplitudes of 2.5 and 5 V the output amplitudes were 0.20 and 0.26 V, respectively. That nonlinear behaviour was less noticeable at tests made at 5 and 10 rad/s.

Figure 10: Control Action $< G_2, \lambda = 120 >$. 

14
5 Conclusions

In this paper, procedures for iterative performance improvement of closed-loop controlled loops are presented with very little prior information on the plant and allowing larger errors at frequencies irrelevant for control design in an attempt to reduce the needed model order, without the need of the model order, poles or zeroes to be similar to that of the elusive “true” plant. Controllers must be robust to the not-yet-excited HF uncertainty, so controller orders can be low only if the plant is expected to have a reasonable downward slope in the frequency response.

The examples show that as performance requirements become more stringent, model complexity usually needs also to do so, but the needed order is not necessary to be known \textit{a priori}, and as plenty of lower-frequency information is available from previous experiments, new ID tests concentrate excitation energy at medium and high frequency bands.

Note that the ideas could be applicable to different approaches: there is ample room for alternatives and discussion in order selection, ID method, model set to choose, uncertainty description, controller design methodology, etc. and the ones presented here are not claimed to be “optimal” under any general criteria.

Acknowledgements

The authors wish to acknowledge the financial support from Spanish Ministerio de Ciencia y Tecnología and their institution. They wish also to thank the reviewers and editors their helpful comments and observations.

References


Figure 13: Best performance achieved with each of the models.


