



MODEL-ROBUST DESIGN OF EXPERIMENTS FOR SEQUENTIAL IDENTIFICATION OF ODE PARAMETERS

Hassan El Abiad, Laurent Le Brusquet, Marie-Ève Davoust

Department of Signal Processing and Electronic Systems, Supélec, Gif-sur-Yvette, France
e-mails : Hassan.ElAbiad, Laurent.LeBrusquet, Marie-Eve.Davoust@supelec.fr

ABSTRACT

This paper presents the idea of sequential model-robust Design of Experiments (DOE) for the identification of dynamic systems modeled with an Ordinary Differential Equation (ODE). The studied DOE problem consists in selecting sequentially the instants where the measures will be done in order to best estimate the system's parameter. The robustness is achieved by considering a statistical representation of the model error defined as the difference between the true ODE and the ODE used in the model. The idea of modeling the model error with a statistical representation has been widely explored in the DOE literature for the identification of static systems. However, there have been little previous works that apply this idea for the identification of dynamic systems. This paper initiates an exploration of this idea in the context of first-order ODE. The model error is modeled by using a kernel-based representation (Gaussian process). A new criterion for the instant selection is constructed and tested on an illustrative example. The design reached with the proposed sequential robust criterion is compared with the design reached with the non-robust version of criterion and with the classical uniform design.

1. INTRODUCTION

For parametric identification problems, the purpose of DOE is to adjust the selection of the experimental conditions to improve the quality of the parameter estimation. This has motivated many researches on DOE over almost a century [1, 2, 3, 4, 5]. The DOE technique was later adapted to the problem of dynamic systems identification [6, 7, 8]. Some of the recent works are summarized in [9].

There have been several works on robust DOE, especially for the identification of static models for engineering design problems. Design robustness with respect to the model error (misspecification) was first discussed by Box and Draper [10] who studied the effect of taking a one-degree polynomial regression model when the target is two-degree. After that, many authors have further discussed and developed this idea with different assumptions about the misspecification [11, 12, 13].

Several recent papers have considered these robust DOE techniques for the identification of ODE parameters. Aspery and Macchietto [14] considered the problem of design robustness with respect to the parameter starting values. They considered that the design is affected by a poor starting values of the parameters and try to derive a criterion that is robust against these starting values. In [15], Goodwin et al. propose a min-max design criterion for the parameter identification of a linear first-order ODE:

$$\dot{x}(t) = a^* x(t) + u(t) + r(t) \quad (1)$$

where $x(t)$ is the state variable, $u(t)$ is the control input, and $r(t)$ is an unknown departure from the simple model $\dot{x}(t) = a^* x(t) + u(t)$ (misspecification). All these functions are scalar and defined on a finite interval $[0, t_{\max}]$. Scalar a^* is the parameter to be identified.

We have studied the same identification problem proposed by Goodwin et al. (equation (1)) but we have considered a statistical approach of the model error $r(t)$ instead of the bound approach derived from the used min-max criterion.

The presence of a non-negligible misspecification has a great influence on the identification of the parameter a^* and also on the selection of the instants t_i where the state variable $x(t)$ will be measured. Misspecified ODE may encounter frequently, especially in the biological field where the main reaction to be modeled is polluted by additional terms whose definition by a specific function would not be feasible.

We define a^* as the value that minimizes the L_2 -norm of the misspecification $r(t)$:

$$a^* = \arg \min_a \int_0^{t_{\max}} (\dot{x}(t) - ax(t) - u(t))^2 dt \quad (2)$$

For a practical reason justified in section 3.1, we choose to estimate $\tau^* = -\frac{1}{a^*}$, the time constant of the system. The aim of designing experiments is to reach an accurate estimation of τ^* with a small number of experiments. This is particularly important when obtaining experiments is difficult, for example for material or cost reasons.

The estimation of τ^* is performed from a database of collected data. Suppose that the initial point ($t_0 = 0, y_0 = x(t_0)$) is known and that a set of n collected data $\{(t_i, y_i) \in]0, t_{\max}] \times \mathbb{R}, i = 1, \dots, n\}$ has already been collected. The t_i 's form the initial design denoted by $\xi_n = [t_1, \dots, t_n]^T$, the y_i 's are noisy observations of $x_i = x(t_i)$. Therefore, the observation equation is defined as:

$$\begin{aligned} y_0 &= x_0 \\ y_i &= x_i + e_i, \quad i = 1 \dots n \end{aligned}$$

where the observations error e_i are normal and i.i.d. ($e_i \sim \mathcal{N}(0, \sigma_e^2)$, σ_e^2 is supposed to be known).

Suppose that we desire to refine the estimation of τ^* by adding a new design point t_{n+1} and its corresponding observation value y_{n+1} to the collected data. Then, the problem of sequential DOE is to choose the next design point t_{n+1} that will best refine the estimation.

An important issue in designing experiments for dynamical system identification is the fact that design criteria generally depend on the unknown parameter which has to be identified. To solve this problem, a sequential DOE strategy is generally implemented. This strategy consists in alternating between parameter estimation stages and designing stages, as shown in figure 1 [2, 16, 17, 18].

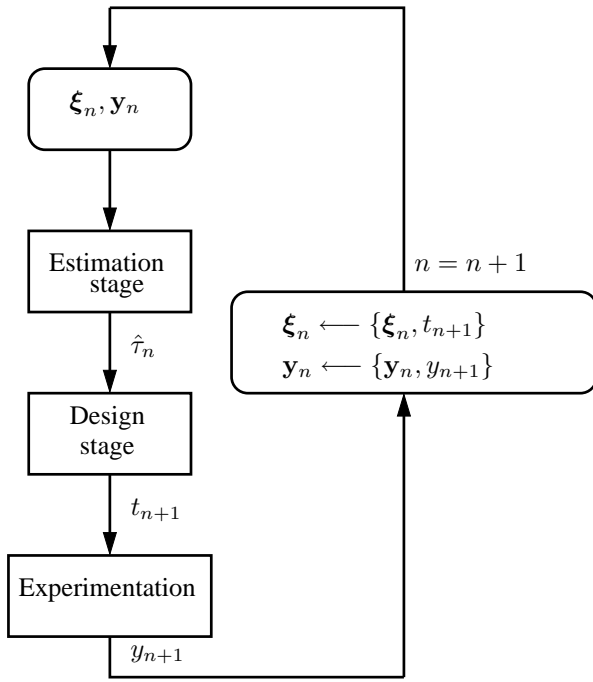


Fig. 1. Sequential DOE

In the estimation stage, the unknown parameter τ^* is estimated using the n collected observations:

$$\hat{\tau}_n = \arg \min_{\tau} J^{\text{estim.}}(\tau, \xi_n, \mathbf{y}_n)$$

In the design stage, the estimated parameter $\hat{\tau}_n$ is used in order to find the next design point t_{n+1} :

$$t_{n+1} = \arg \min_{t > t_n} J^{\text{design}}(t, \hat{\tau}_n)$$

The paper is organized as follows. The next section illustrates the proposed approach and in particular the statistical representation used for $r(t)$ and the derived design criterion $J^{\text{design}}(t, \hat{\tau}_n)$. Mathematical developments for calculating the criteria $J^{\text{estim.}}(\tau, \xi_n, \mathbf{y}_n)$ and $J^{\text{design}}(t, \hat{\tau}_n)$ are not straightforward. They are detailed in section 3. In the fourth section, an example is presented in order to illustrate the improvements of the proposed sequential design over the classical uniform design and the non-robust version of the proposed sequential design. Section five concludes the paper by presenting the main contributions.

2. PROPOSED APPROACH SCHEME

This section presents the proposed sequential robust criterion $J^{\text{design}}(t, \hat{\tau}_n)$. As mentioned in the introduction, the main contribution of this work is the presence of the model error $r(t)$ which improves the robustness of our design criterion.

The model error $r(t)$ is generally a relative smooth function. This is the reason why we model it by a Gaussian process as in [19, 20] (see [21] for a Gaussian process tutorial). A Gaussian process is a random field defined by its mean and covariance function:

$$\begin{aligned} \mathbb{E}_r[r(t)] &= 0, & \forall t \in [0, t_{\max}] \\ \mathbb{E}_r[r(t)r(t')] &= c(t - t'), & \forall (t, t') \in [0, t_{\max}]^2 \end{aligned}$$

where $\mathbb{E}[\cdot]$ denotes the expected value and $s^2 = c(0)$ denotes the variance of the model error. The relevance of modeling the misspecification by a Gaussian process rises because for some classes of covariance functions, Gaussian processes span a rather large space (infinite-dimensional). Therefore, this type of representation matches the robustness requirement: the design point t_{n+1} that we look for must lead to a good estimation performance whatever is the misspecification. The misspecification is unknown but our idea of robustness states that the chosen design point must guarantee a good level of performance (on average) over a wide range of potential misspecifications. The theoretical expression of the proposed criterion is directly derived from the use of the statistical representation of $r(t)$ that was suggested:

$$J^{\text{design}}(t, \hat{\tau}_n) = \mathbb{E}_{\{e_n, e_{n+1}, r\}} [(\hat{\tau}_{n+1}(t) - \tau^*)^2] \quad (3)$$

where $\hat{\tau}_{n+1}$ is the updated estimation of τ^* and $\mathbf{e}_n = [e_1, \dots, e_n]^T$. The use of the expected value over r (i.e. $r(t), t \in [0, t_{\max}]$)

assures the robustness with respect to the model misspecification.

3. MATHEMATICAL DEVELOPMENTS OF THE CRITERIA

This section presents the mathematical developments of the estimation and design criteria introduced in the previous section.

3.1. Estimation stage criterion

The first stage of the DOE procedure consists in estimating the unknown parameter τ^* given n observations. This section presents a detailed formulation of the parameter estimation procedure. Two estimators of τ^* are presented: the proposed robust estimator (considering model-error $r(t)$ and measurement error $e(t)$) and its non-robust version (considering only measurement error $e(t)$). The non-robust version will be used for comparison in section 4.

First, let's use a numerical integration scheme in order to discretize the state equation (1) at the instants t_i :

$$x(t_i) = x_0 + a^* \int_0^{t_i} x(\theta) d\theta + \int_0^{t_i} u(\theta) d\theta + \int_0^{t_i} r(\theta) d\theta \quad (4)$$

where $i = 1, \dots, n$. A numerical integration scheme is preferred over a derivative one for stability reasons.

The second term in (4) is approximated by the trapezoidal integration method. Thus, $\mathbf{x}_n = [x(t_1), \dots, x(t_n)]^\top$ can be written in vector form:

$$\mathbf{x}_n = x_0 + a^* \mathbf{W}_n \mathbf{x}_n + \mathbf{u}_n + \mathbf{r}_n \quad (5)$$

where $\mathbf{u}_n = [\int_0^{t_1} u(\theta) d\theta, \dots, \int_0^{t_n} u(\theta) d\theta]^\top$ is the control input vector defined by the input signal $u(t)$ which is known.

$\mathbf{r}_n = [\int_0^{t_1} r(\theta) d\theta, \dots, \int_0^{t_n} r(\theta) d\theta]^\top$ is unknown and its possible values depend on the used statistical representation. The $n \times n$ matrix \mathbf{W}_n contains the coefficients of the trapezoidal integration method.

The estimation of τ^* is performed from the collected data $\mathbf{y}_n = [y(t_1), \dots, y(t_n)]^\top$ which are noisy observations of (5):

$$\mathbf{y}_n = x_0 + a^* \mathbf{W}_n (\mathbf{y}_n - \mathbf{e}_n) + \mathbf{u}_n + \mathbf{e}_n + \mathbf{r}_n$$

The estimator $\hat{\tau}_n$ must be coherent with the definition of τ^* . In (2), τ^* is obtained by minimizing the L_2 -norm, which is a natural Maximum Likelihood criterion when errors are i.i.d. The same assumptions are kept to construct the estimator $\hat{\tau}_n$. Thus, the obtained estimator will be consistent ($\hat{\tau}_n \rightarrow \tau^*$), which is the main importance for an estimator. We will see in section 3.2 that a more realistic

covariance function will be used to compute a precise value of $\mathbb{E}_{\{\mathbf{e}_n, r\}} [(\hat{\tau}_n - \tau^*)^2]$ (main importance for the evaluation of the quality of an estimator).

The $\hat{\tau}_n$ estimator is thus obtained by considering the Maximum Likelihood of the vector $\mathbf{e}_n + \mathbf{r}_n$ with $r(t)$ supposed to be i.i.d. ($r(t) \sim \mathcal{N}(0, s^2)$). When supposing $r(t)$ i.i.d., the vector \mathbf{r}_n is a Gaussian vector. If $k \leq j$, the (k, j) coefficient of its $n \times n$ covariance matrix \mathbf{D}_n is equal to $s^2 t_k$. Thus, the Likelihood criterion is constructed with a covariance matrix equal to $(\mathbf{D}_n + \sigma_e^2 \mathbf{I}_n)$ with \mathbf{I}_n a $n \times n$ identity matrix. Therefore, The estimator expression is given by:

$$\hat{\tau}_n = \frac{\mathbf{y}_n^\top \mathbf{W}_n^\top (\mathbf{D}_n + \sigma_e^2 \mathbf{I}_n)^{-1} \mathbf{W}_n \mathbf{y}_n}{\mathbf{y}_n^\top \mathbf{W}_n^\top (\mathbf{D}_n + \sigma_e^2 \mathbf{I}_n)^{-1} [\mathbf{y}_n - \mathbf{v}_n]^\top} \quad (6)$$

where $\mathbf{v}_n = x_0 + \mathbf{u}_n$.

The non-robust estimator is simply obtained by considering $\mathbf{D}_n = 0$ in the previous equation.

3.2. Design stage criterion

The expectation computation in (3) may be reached because, for a given misspecification r , a given design, and a given noise realization $[e_n, e_{n+1}]$, it is possible to write analytically the expression of $\hat{\tau}_{n+1}$. Indeed, the solution of (1) is:

$$x(t) = x_0 \exp(a^* t) + \exp(a^* t) \int_0^{t_{\max}} (r(\theta) + u(\theta)) \exp(-a^* \theta) d\theta \quad (7)$$

The integral terms in the previous equation are obtained by a fine discretization of the experimental domain $[0, t_{\max}]$. Let $\{t'_i\}_{i=1 \dots m}$ be this time discretization. In the following, subscript ($'$) denotes vector calculated at these instants (e.g. $\mathbf{r}'_m = [r(t'_1) \dots r(t'_m)]^\top$).

Thus, equation (7) can be written in vector form:

$$\mathbf{x}'_m = \mathbf{P}'_m (\mathbf{r}'_m + \mathbf{u}') + \mathbf{f}'_m$$

where \mathbf{P}'_m is the $m \times m$ matrix of integration coefficients (classical trapezoidal integration including exponential function) and $\mathbf{f}'_m = [x_0 \exp(a^* t'_1) \dots x_0 \exp(a^* t'_m)]^\top$ is the function $f(t) = x_0 \exp(a^* t)$ evaluated at the m discretization instants.

Following the same procedure, the n collected data can be written:

$$\mathbf{y}_n = \mathbf{P}_n (\mathbf{r}'_m + \mathbf{u}'_m) + \mathbf{f}_n + \mathbf{e}_n \quad (8)$$

where \mathbf{P}_n is a $n \times m$ matrix containing the integration coefficients for the n collected data, and \mathbf{f}_n is the function $f(t)$ evaluated at observation instants (t_1, \dots, t_n) .

The sequential DOE consists in choosing the next design point t_{n+1} which will refine the parameter estimation.

To estimate the unknown parameter τ^* using $n + 1$ observations, we have to adopt the estimation expression in (6):

$$\hat{\tau}_{n+1} = \frac{\mathbf{y}_{n+1}^\top \mathbf{W}_{n+1}^\top (\mathbf{D}_{n+1} + \sigma_e^2 \mathbf{I}_{n+1})^{-1} \mathbf{W}_{n+1} \mathbf{y}_{n+1}}{\mathbf{y}_{n+1}^\top \mathbf{W}_{n+1}^\top (\mathbf{D}_{n+1} + \sigma_e^2 \mathbf{I}_{n+1})^{-1} [\mathbf{y}_{n+1} - \mathbf{v}_{n+1}]^\top} \quad (9)$$

Using (8), the expression of the parameter estimation (9) is rewritten in the following form:

$$\hat{\tau}_{n+1} = \frac{\mathbf{r}'_m{}^\top \mathbf{M}_{\text{num}} \mathbf{r}'_m + \mathbf{r}'_m{}^\top \mathbf{v}_{\text{num}} + c_{\text{num}}}{\mathbf{r}'_m{}^\top \mathbf{M}_{\text{den}} \mathbf{r}'_m + \mathbf{r}'_m{}^\top \mathbf{v}_{\text{den}} + c_{\text{den}}} = \frac{\text{Num}}{\text{Den}} \quad (10)$$

where,

$$\begin{aligned} \mathbf{M}_{\text{den}} &= \mathbf{P}_{n+1}^\top \mathbf{W}_{n+1}^\top (\mathbf{D}_{n+1} + \sigma_e^2 \mathbf{I}_{n+1})^{-1} \mathbf{P}_{n+1} \\ \mathbf{v}_{\text{den}} &= \mathbf{P}_{n+1}^\top \mathbf{W}_{n+1}^\top (\mathbf{D}_{n+1} + \sigma_e^2 \mathbf{I}_{n+1})^{-1} \\ &\quad (\mathbf{f}_{n+1} + \mathbf{e}_{n+1} - \mathbf{x}_0) \\ &\quad + \mathbf{P}_{n+1}^\top (\mathbf{D}_{n+1} + \sigma_e^2 \mathbf{I}_{n+1})^{-1} \mathbf{W}_{n+1} (\mathbf{f}_{n+1} + \mathbf{e}_{n+1}) \\ c_{\text{den}} &= (\mathbf{f}_{n+1} + \mathbf{e}_{n+1} - \mathbf{x}_0)^\top (\mathbf{D}_{n+1} + \sigma_e^2 \mathbf{I}_{n+1})^{-1} \\ &\quad \mathbf{W}_{n+1} (\mathbf{f}_{n+1} + \mathbf{e}_{n+1}) \\ \mathbf{M}_{\text{num}} &= \mathbf{P}_{n+1}^\top \mathbf{W}_{n+1}^\top (\mathbf{D}_{n+1} + \sigma_e^2 \mathbf{I}_{n+1})^{-1} \mathbf{W}_{n+1} \mathbf{P}_{n+1} \\ \mathbf{v}_{\text{num}} &= \mathbf{P}_{n+1}^\top \mathbf{W}_{n+1}^\top ((\mathbf{D}_{n+1} + \sigma_e^2 \mathbf{I}_{n+1})^{-1})^{-1} \\ &\quad + [(\mathbf{D}_{n+1} + \sigma_e^2 \mathbf{I}_{n+1})^{-1}]^\top \\ &\quad \mathbf{W}_{n+1}^\top (\mathbf{f}_{n+1} + \mathbf{e}_{n+1}) \\ c_{\text{num}} &= \mathbf{W}_{n+1}^\top (\mathbf{f}_{n+1} + \mathbf{e}_{n+1})^\top (\mathbf{D}_{n+1} + \sigma_e^2 \mathbf{I}_{n+1})^{-1} \\ &\quad \mathbf{W}_{n+1} (\mathbf{f}_{n+1} + \mathbf{e}_{n+1}) \end{aligned}$$

and \mathbf{x}_0 is the $n + 1$ vector $[x_0, \dots, x_0]^\top$.

As the parameter estimation expression (10) is derived, the criteria derivation is accomplished by taking the expectation of $(\hat{\tau}_{n+1} - \tau^*)^2$ with respect to the model and observation errors as shown in (3). At this step, as τ^* is unknown, it is replaced by its estimation $\hat{\tau}_n$. This is the purpose of the two-stage sequential DOE presented in the introduction. Expression (10) can be written as follows:

$$\hat{\tau}_{n+1} - \tau^* \approx \frac{\mathbf{r}'_m{}^\top \mathbf{M} \mathbf{r}'_m + \mathbf{r}'_m{}^\top \mathbf{v} + c}{\mathbf{r}'_m{}^\top \mathbf{M}_{\text{den}} \mathbf{r}'_m + \mathbf{r}'_m{}^\top \mathbf{v}_{\text{den}} + c_{\text{den}}} = \frac{A}{\text{Den}} \quad (11)$$

where, $\mathbf{M} = \mathbf{M}_{\text{num}} - \hat{\tau}_n \mathbf{M}_{\text{den}}$, $\mathbf{v} = \mathbf{v}_{\text{num}} - \hat{\tau}_n \mathbf{v}_{\text{den}}$ and $c = c_{\text{num}} - \hat{\tau}_n c_{\text{den}}$. Doing statistics on (11) is not evident because of the presence of stochastic variables in the numerator and denominator. Fortunately, for reasonable choice of the model error variance s^2 , the variations of $\epsilon_{\text{den}} = \mathbf{r}'_m{}^\top \mathbf{M}_{\text{den}} \mathbf{r}'_m + \mathbf{r}'_m{}^\top \mathbf{v}_{\text{den}}$ are negligible compared to c_{den} :

$$\frac{A}{\text{Den}} = \frac{\epsilon_A + c}{\epsilon_{\text{den}} + c_{\text{den}}} \approx \frac{c}{c_{\text{den}}} + \frac{1}{c_{\text{den}}} \epsilon_A - \frac{c}{c_{\text{den}}^2} \epsilon_{\text{den}}$$

Furthermore, it was noticed that the variations of $\frac{c}{c_{\text{den}}^2} \epsilon_{\text{den}}$ are negligible compared to the variations of $\frac{1}{c_{\text{den}}} \epsilon_A$. Therefore, the denominator is supposed to be constant (equal to

c_{den}). This point is illustrated in the next section. Now, doing statistics on (11) is possible because Den is constant and one has to do statistics only on A . This justifies our suggestion of working with τ^* instead of a^* .

Knowing that Den is constant, the criterion in (3) is approximated by:

$$J^{\text{design}}(t, \hat{\tau}_n) = \frac{1}{\text{Den}} \mathbb{E}_{\{\mathbf{e}_n, \mathbf{e}_{n+1}, r\}} [A^2]$$

The stochastic variable $A = \mathbf{r}'_m{}^\top \mathbf{M} \mathbf{r}'_m + \mathbf{r}'_m{}^\top \mathbf{v} + c$ consists of the products of correlated Gaussian variables. Therefore, it may be computed thanks to a singular value decomposition:

$$A = \mathbf{z}^\top \mathbf{S}_z \mathbf{z} + \mathbf{z}^\top \mathbf{v}_z + c_z$$

where c_z is a constant value, $\mathbf{z} \sim \mathcal{N}(0, \mathbf{I}_{n+1})$, and \mathbf{S}_z is a diagonal matrix.

The mean m_A and the variance σ_A^2 of A can be calculated as functions of \mathbf{S}_z , \mathbf{v}_z and c_z :

$$\begin{aligned} m_A &= \sum_k \lambda_k + c_z \\ \sigma_A^2 &= 2 \sum_k \lambda_k^2 + \left(\sum_k \lambda_k \right)^2 + \sum_k v_k^2 \end{aligned}$$

where λ_k is the k -th element in the diagonal of \mathbf{S}_z and v_k is the k -th element of \mathbf{v}_z . The expression of J^{design} is thus tractable:

$$J^{\text{design}}(t, \hat{\tau}_n) = \frac{1}{\text{Den}} (m_A^2 + \sigma_A^2)$$

4. ILLUSTRATIVE EXAMPLE

In this section, an illustrative example is given in order to show the advantage of performing DOE (comparison of designs reached by a sequential DOE approach and other designs reached with the classical uniform criterion) and the advantage of considering a model error (comparison of the proposed design with the non-robust design obtained by taking $\mathbf{D}_n = 0$ in the estimation stages and $\mathbf{r}_n = 0$ in the design stages).

In this example, the control input is taken to be $u(t) = 0$ and the system is only allowed for the initial condition $x_0 = 1$.

The Gaussian kernel is used because it is the most used kernel for the Gaussian process covariance [19]:

$$c(t - t') = s^2 \exp \left[- \left(\frac{t - t'}{\lambda} \right)^2 \right], \quad \forall (t, t') \in [0, t_{\text{max}}] \quad (12)$$

where, s^2 (Gaussian process variance) and λ (correlation distance) are the Gaussian process parameters. The kernel is used with the Gaussian process parameters values of $s^2 = 0.01$ and $\lambda = 0.2$. The approach used to choose these values is based on a maximin efficiency criterion (see [19, 20] for more details).

Having a set of initial points between the initial time value ($t_0 = 0$) and the intermediate time value (t_{int}), the goal is to select the time values between t_{int} and t_{max} in order to best estimate the parameter τ^* . Let $t_{\text{max}} = 30$ s the final time, $t_{\text{int}} = 1$ s and $\xi_3 = [0.33, 0.66, 1]^\top$ the initial design vector. The initial parameter value $a^* = -0.5$ and the time constant corresponding value is $\tau^* = 2$ s.

The proposed approach has been tested with a Monte-Carlo method where 200 identification problems were created. Each problem is defined by a particular realization of the Gaussian process $r(t)$ and by a particular realization of the observation error e_n with $\sigma_e^2 = 0.05$.

In section 3, an assumption about Den was made. We have assumed that the variations of $\mathbf{r}'_m^\top \mathbf{M}_{\text{den}} \mathbf{r}'_m + \mathbf{r}'_m^\top \mathbf{v}_{\text{den}}$ are negligible in front of the constant term c_{den} . Therefore, Den was considered to be constant $Den \approx c_{\text{den}}$. Figure 2 gives the histograms of Den and $\epsilon_A \times \frac{c_{\text{den}}}{c}$ obtained from the 200 realizations of $r(t)$. The variations of Den values are small enough compared to c_{den} and to the variations of $\epsilon_A \times \frac{c_{\text{den}}}{c}$. Therefore, the assumption about Den and the notice about the variations of $\frac{c}{c_{\text{den}}} \epsilon_{\text{den}}$ made on section 3 are valid.

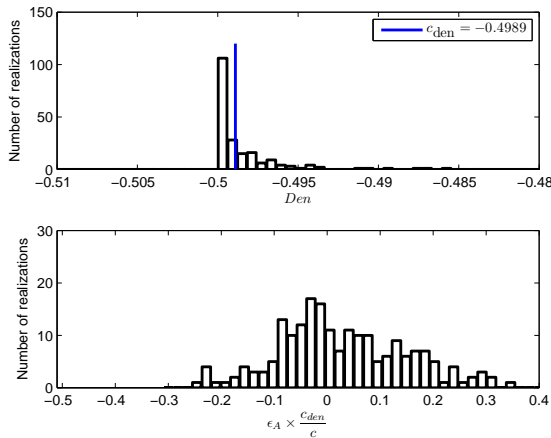


Fig. 2. Histograms of Den and $\epsilon_A \times \frac{c_{\text{den}}}{c}$

Figure 3 shows the mean of the square errors ($\hat{\tau}_{n+1} - \tau^*$)² when 30 design points are added. The comparison is done among the proposed sequential approach, the non-robust version of the sequential approach and the uniform approach where the design points are distributed uniformly over the experimental domain. The results show a faster

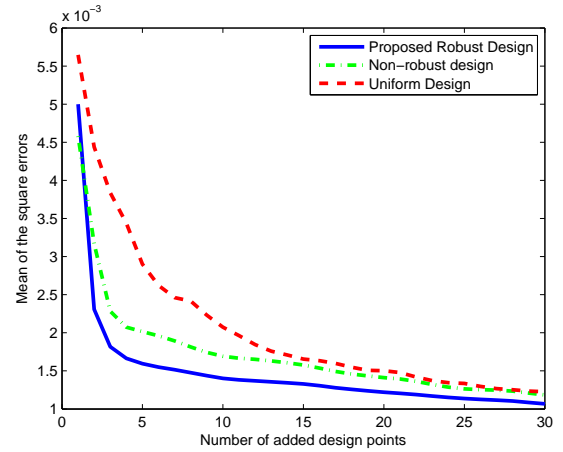


Fig. 3. Comparison among the three criteria

convergence of the proposed approach over the other two approaches.

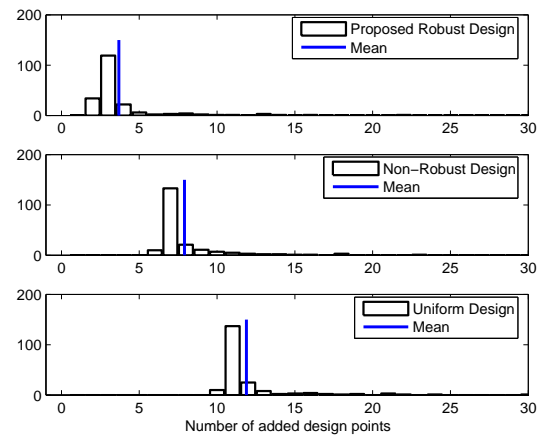


Fig. 4. Histograms of the number of added points with a threshold equal to 1.8×10^{-3}

Another way of comparison is to compare the number of added design points to reach a square error lower than a fixed threshold. Figure 4 compares the histograms of the added design points for the three criteria. The corresponding means are shown in Table 1. It is clear that the proposed approach is the most efficient because it needs the smallest number of added points. This result may be particularly interesting when the number of experiments has to be small.

	Robust	Non-robust	Uniform
Mean Added Points	3.69	7.90	11.88

Table 1. Means of the number of added design points

5. CONCLUSION AND FUTURE WORK

This paper has discussed the idea of designing robust experiments for dynamic systems. The main contribution of this paper is the idea of modeling the misspecification by a statistical representation (Gaussian process). The paper has presented a detailed derivation of the proposed criterion showing the simplifications and main contributions. Finally, an illustrative example was presented showing the improvement in the speed of parameter convergence over other criteria.

The proposed idea is tested on a very simple context since it is restricted to first-order linear ODE. The goal of the paper is to show the relevance of using a statistical representation associated with a statistical criterion in order to reach robust estimators. Further works on more complex systems (non-scalar and non-linear ODE) are in progress. We also plan to exploit the proposed criterion to design for both t_n and the value of the control input.

6. REFERENCES

- [1] A. Wald, "On the efficient design of statistical investigations," *Annals of Mathematical Statistics*, vol. 14, pp. 134–140, 1943.
- [2] J. Wolfowitz J. Kiefer, "The equivalence of two extremum problems," *Canadian Journal of Mathematics*, vol. 12, pp. 363–366, 1960.
- [3] H. P. Wynn, "Results in the theory and construction of d -optimum experimental designs," *Journal of Royal Statistical Society*, vol. 2, pp. 133–147, 1972.
- [4] V. V. Federov, *Theory of optimal experiments*, Academic Press, New York, London, 1972.
- [5] P. Whittle, "Some general points in the theory of optimal experimental design," *Journal of Royal Statistical Society*, vol. 1, pp. 123–130, 1973.
- [6] V. S. Levadi, "Design of input signals for parameter estimation," *IEEE Transactions on Automatic Control*, vol. 2, pp. 205–211, 1966.
- [7] S. Arimoto and H. Kimura, "Optimal input test signals for system identification - an information theoretic approach," *Int. J. Systems Sci.*, vol. 2, pp. 279–290, 1973.
- [8] R. K. Mehra, "Optimal inputs for system identification," *IEEE Transactions on Automatic Control*, p. 192–200, 1974.
- [9] M. Gevers, "Identification for control: From the early achievements to the revival of experiment design," *European Journal of Control*, vol. 11, pp. 1–18, 2005.
- [10] G. E. P. Box and N. R. Draper, "A basis for the selection of a response surface design," *Journal of the American Statistical Association*, vol. 54, pp. 622–654, 1959.
- [11] G. Montepiedra and V. V. Fedorov, "Minimum bias designs with constraints," *Journal of Statistical Planning and Inference*, vol. 63, pp. 97–111, 1997.
- [12] R. X. Yue and F. J. Hickernell, "Robust designs for fitting linear models with misspecification," *Statistica Sinica*, vol. 9, pp. 1053–1069, 1999.
- [13] Z. Fang and D. P. Wiens, "Robust regression designs for approximate polynomial models," *Journal of Statistical Planning and Inference*, vol. 117, pp. 305–321, 2003.
- [14] S.P. Asprey and S. Macchietto, "Designing robust optimal dynamic experiments," *Journal of process control*, vol. 12, pp. 545–556, 2002.
- [15] G. C. Goodwin C. R. Rojas, J. S. Welsh and A. Feuer, "Robust optimal experiment design for system identification," *Automatica*, vol. 43, pp. 993–1008, 2007.
- [16] C. F. J. Wu, "Asymptotic inference from sequential design in a nonlinear situation," *Biometrika*, vol. 3, pp. 553–558, 1985.
- [17] C. Kulcsar L. Pronzato and E. Walter, "The interface between sequential experimental design and dual control," *Statist. Sinica*, vol. 5, pp. 421–440, 1995.
- [18] L. Pronzato E. Walter, *Identification of Parametric Models from Experimental Data*, Springer, Communications and Control Engineering Series, Londres, 1997.
- [19] M. Roger, L. Le Brusquet, and G. Fleury, "A criterion for model-robust design of experiments," in *IEEE International Workshop on Machine Learning for Signal Processing*, Sao Luis (Brazil), September 29–October 1st, 2004, pp. 33–42.
- [20] H. El Abiad, L. Le Brusquet, M. Roger, and M. Davoust, "Model-Robust Sequential Design of Experiments for Identification Problems," in *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Honolulu, Hawaii (US), April 15–April 20, 2007, vol. 2, pp. 441–444.
- [21] C. K. I. Williams, "Prediction with Gaussian processes: From linear regression to linear prediction and beyond," Tech. Rep. NCRG/97/012, Department of Computer Science and Applied Mathematics, Aston University, Birmingham, UK, Oct. 1997.