This document contains the **post-print pdf-version** of the refereed paper:

“Optimal experiment design for dynamic bioprocesses: a multi-objective approach”

by *Dries Telen, Filip Logist, Eva Van Derlinden, Ignace Tack and Jan Van Impe*

which has been archived in the university repository Lirias (<https://lirias.kuleuven.be/>) of the Katholieke Universiteit Leuven.

The content is identical to the content of the published paper, but without the final typesetting by the publisher.

When referring to this work, please cite the full bibliographic info:

Telen, D., Logist, F., Van Derlinden, E., Tack, I., Van Impe, J., 2012. Optimal experiment design for dynamic bioprocesses: a multi-objective approach. Chemical Engineering Science 78, 82–97.

The journal and the original published paper can be found at:

<http://www.sciencedirect.com/science/journal/00092509>
<http://dx.doi.org/10.1016/j.ces.2012.05.002>

The corresponding author can be contacted for additional info.

Conditions for open access are available at:

<http://www.sherpa.ac.uk/romeo/>

Optimal experiment design for dynamic bioprocesses: a multi-objective approach

D. Telen^a, F. Logist^a, E. Van Derlinden^a, I. Tack^a, J. Van Impe^{a,*}

^a*BioTeC & OPTEC, Chemical Engineering Department, Katholieke Universiteit Leuven,
W. de Croylaan 46, 3001 Leuven, Belgium*

Abstract

Dynamic process models can be used for operating, controlling and optimising important bioprocesses, e.g., pharmaceuticals production, enzyme production and brewing. After selecting an appropriate process model structure, parameter estimates have to be obtained based on real-life experiments. To reduce the amount of labour and often cost intensive experiments optimal experiment design (OED) is an indispensable tool. In optimal experiment design, dynamic input profiles have to be determined in order to obtain informative experiments. In the particular case of optimal experiment design for parameter estimation, a scalar measure of the Fisher Information Matrix is used as an objective function. Over the years different criteria have been developed. However, the important question that remains is *which criterion to choose*. In this work, an approach to tackle the criterion selection is presented. In addition, a multi-objective optimisation approach is implemented, which enables to combine two, often competing optimisation criteria. The developed approach is illustrated with two case studies. The first case study is a fed-batch bioreactor model and the second case study is a Lotka Volterra fishing model.

Keywords: Optimal experiment design, Multi-objective optimisation, Parameter estimation, Dynamic optimisation, Bioprocess

1. Introduction

An important tool for the operation, control and optimisation of (bio)processes is the mathematical process model. After a suitable model structure is iteratively determined, accurate parameter values have to be estimated as the next step in the modelling cycle (Ljung, 1999; Walter and Pronzato, 1997; Franceschini and Macchietto, 2008a). Performing experiments can, however, be cost and labour intensive. Optimal experiment design (OED) techniques are appealing to limit the experimental burden while at the same time maximising the information content in the experiments (Balsa-Canto et al., 2008; Van Derlinden et al., 2010; Donckels et al., 2010; Kud et al., 2010). As processes in the (bio)chemical field typically exhibit a dynamic nature, they are described by ordinary differential equations. Consequently, the time varying inputs to these models can be optimised (Bauer et al., 2000; Balsa-Canto et al., 2001) in order to obtain informative experiments. Traditionally, the information content of an experiment is quantified using the Fisher Information Matrix (Fisher, 1935) and a specific scalar function of this Fisher Information Matrix is used as the objective function.

*Corresponding author

Email address: jan.vanimpe@cit.kuleuven.be (J. Van Impe)

Although several different criteria have been proposed (Walter and Pronzato, 1997; Franceschini and Macchietto, 2008a,b), the question remains which *criterion to choose* in practice as each of these criteria has its strengths and weaknesses. In addition, some of the criteria are even conflicting. Possible trade-offs between criteria are, e.g., maximising parameter accuracy and minimising the correlation between parameters. The multi-objective optimisation approach, as presented by Logist et al. (2011b) offers a method to study the trade-offs between the different criteria and mitigate their drawbacks. In the paper by Logist et al. (2011b) the feasibility of formulating optimal experiment design problems in a multi-objective framework within the ACADO environment was presented. Sometimes this multi-objective character of a problem is not acknowledged as such. For example, in Franceschini and Macchietto (2008b) several novel decorrelation criteria are introduced. In practice these criteria boil down to a parameter accuracy criterion and a decorrelation criterion. Typically, one is constrained to a predetermined value while the other is optimised and vice versa. It has to be noted that this approach mimics the ϵ -constraint method for multi-objective optimisation described by Haimes et al. (1971). However, only a single solution out of the set of trade-off or also called Pareto optimal solutions is returned and no insight is provided about trade-offs between different alternatives. As this Fisher Information Matrix depends on the initial guess of the parameters, sometimes a more robust optimal experiment is required (Körkel et al., 2004). This robustness aspect can then be considered as an additional objective (Logist et al., 2011a). In this latter work, robustness against parametric uncertainty has been taken into account in optimal control problems.

The current study alleviates the need for a-priori specification of these constraint values as it systematically provides the entire Pareto set. Afterwards a sound selection of the experiments to be performed can be made by the decision maker, e.g., the process engineer.

This paper is structured as follows. In Section 2, the mathematical formulation of a dynamic, (multi-objective) optimal experiment design is presented. To conclude Section 2 the solution strategy is described. Section 3 introduces the two employed case studies. The obtained results for the two case studies are discussed in Section 4.

2. Mathematical formulation

This section is structured as follows. First, the problem class of dynamic optimisation is addressed, as optimal experiment design is a subclass of the more general class optimal control problems. In the following subsection, the specific objective functions with respect to optimal experiment design are introduced. The third subsection explains how the multi-objective optimisation aspect is dealt with in this work as the trade-off between the different OED objectives is framed within a multi-objective context. The proposed strategy on how to choose a criteria is discussed in the fourth subsection. The used software and its implementation conclude this section.

2.1. Dynamic optimisation

The general formulation of dynamic optimisation reads as follows:

$$\min_{\mathbf{x}(\xi), \mathbf{u}(\xi), \mathbf{p}, \xi} J(\mathbf{x}(\xi), \mathbf{u}(\xi), \mathbf{p}, \xi) \quad (1)$$

subject to:

$$\frac{d\mathbf{x}}{d\xi} = \mathbf{f}(\mathbf{x}(\xi), \mathbf{u}(\xi), \mathbf{p}, \xi) \quad (2)$$

$$0 = \mathbf{b}_c(\mathbf{x}(0)) \quad (3)$$

$$0 \geq \mathbf{c}_p(\mathbf{x}(\xi), \mathbf{u}(\xi), \mathbf{p}, \xi) \quad (4)$$

$$0 \geq \mathbf{c}_t(\mathbf{x}(\xi_f), \mathbf{u}(\xi_f), \mathbf{p}, \xi_f) \quad (5)$$

where $\xi \in [0, \xi_f]$, $\mathbf{x}$ are the state variables, $\mathbf{u}$ are the control variables and $\mathbf{p}$ are the parameters. The dynamic system equations are represented by $\mathbf{f}$. The vector $\mathbf{b}_c$ indicates the initial conditions, the vector $\mathbf{c}_p$, the path inequality constraints and $\mathbf{c}_t$ the terminal inequality constraints on controls and states. The design freedoms can be grouped together as $\mathbf{z} = [\mathbf{x}(\xi), \mathbf{u}(\xi), \xi_f]$. All vectors $\mathbf{z}$ that satisfy equations (2) to (5) form the set of feasible solutions S . This optimisation problem is infinite dimensional. These problems are nowadays solved by converting them to a finite Non-Linear Programming (NLP) problem by means of discretisation. There are two different approaches. A *sequential direct method* as *Single Shooting* (e.g., Sargent and Sullivan (1978); Vassiliadis et al. (1994a,b)) discretises the controls while the *simultaneous direct approaches* discretise both states and controls. Within these simultaneous approaches there exists: *Multiple Shooting* (e.g., Bock and Plitt (1984); Leineweber et al. (2003)) and *Orthogonal Collocation* (e.g., Biegler (1984, 2007)).

2.2. Optimal experiment design for parameter estimation

In this paper the optimal experiment design for parameter estimation is considered. In order to quantify the amount of information for parameter estimation in an experiment the Fisher Information Matrix approach is employed (Walter and Pronzato, 1997). The Fisher Information Matrix is defined as:

$$\mathbf{F}(\mathbf{p}) = \int_0^{t_f} \left(\frac{\partial \mathbf{y}(\mathbf{p}, t)}{\partial \mathbf{p}} \right)^T \bigg|_{\mathbf{p}=\mathbf{p}^*} \mathbf{Q} \left(\frac{\partial \mathbf{y}(\mathbf{p}, t)}{\partial \mathbf{p}} \right) \bigg|_{\mathbf{p}=\mathbf{p}^*} dt \quad (6)$$

As the true values $\mathbf{p}^*$ are unknown, the Fisher Information Matrix is evaluated at the current best guess. The Fisher Information Matrix combines information about the output error measurements, $\mathbf{Q}$ and the sensitivities of the output model with respect to small variations in the model parameters, $\frac{\partial \mathbf{y}(\mathbf{p}, t)}{\partial \mathbf{p}}$. An interesting property of the Fisher Information Matrix is that under the assumption of unbiased estimators and uncorrelated Gaussian noise, the inverse of $\mathbf{F}$ approximates the lower bound of the parameter estimation variance-covariance matrix, which is the Cramér-Rao lower bound (Ljung, 1999).

The matrix $\mathbf{F}$ contains the information of a chosen experiment. To optimise this amount of information a scalar function of $\mathbf{F}$ has to be selected as the optimisation problem's objective function. Mathematically expressed:

$$J = \Phi(\mathbf{F}(\mathbf{x}(\xi), \mathbf{u}(\xi), \mathbf{p}, \xi)) \quad (7)$$

where $\Phi(\cdot)$ is one of the criteria further described and $\mathbf{F}$ is the Fisher Information Matrix. This illustrates that optimal experiment design problems form a specific subclass of dynamic optimisation. The scalar functions considered in this work are listed below:

A-criterion: $\min[\text{trace}(\mathbf{F}^{-1})]$ A-optimal designs minimise the mean of the parameter estimation errors. The geometrical interpretation is the minimisation of

the enclosing frame around the joint confidence region. In order to decrease the computational effort, the problem is often reformulated as maximising the trace of $\mathbf{F}$.

D-criterion: $\max[\det(\mathbf{F})]$ D-optimal designs minimise the geometric mean. Geometrically, this is minimising the volume of the joint confidence region.

E-criterion: $\max[\lambda_{\min}(\mathbf{F})]$ E-optimal designs aim at minimising the largest parameter error. Uncertainty regarding other parameters is thus neglected. This corresponds to minimising the length of the largest uncertainty axis of the joint confidence region.

Modified E-criterion: $\min[\frac{\lambda_{\max}(\mathbf{F})}{\lambda_{\min}(\mathbf{F})}]$ The modified E-criterion (ME-criterion) minimises the condition number of the Fisher Information Matrix. A priori the optimal value of one is known. This corresponds to circular joint confidence regions.

M-criterion: $\max[\min \mathbf{F}_{ii}]$ The M-criterion tries to minimise directly the uncertainty of the most uncertain parameter. This criterion is very analogous to the E-criterion as both work directly on the most uncertain parameter.

The aforementioned criteria are widely known and used (Pukelsheim, 1993; Walter and Pronzato, 1997; Franceschini and Macchietto, 2008a).

Recently several new decorrelation criteria are developed by Franceschini and Macchietto (2008b). The basic idea is to maximise information content while at the same time satisfying some decorrelation threshold or vice versa. The general formulation is:

$$\min h(\mathbf{C}) \quad (8)$$

which means minimising over some function, usually a sum, of the correlation coefficients. These new criteria can however also be considered within a multi-objective frame, e.g., trying to maximise the information content and at the same time to minimise the correlation.

Each of the described criteria has also some known drawbacks (Franceschini and Macchietto, 2008a; Walter and Pronzato, 1997).

Given these reported drawbacks, it is clear that there is no single optimal criterion, so the important question of which criterion to choose remains. Moreover, some of these criteria can even be conflicting. This means that the combination of criteria can be a solution to some of the reported problems. Similar ideas have been presented in e.g., Versyck and Van Impe (1998); Alberton et al. (2010). However the current paper casts this issue in a multi-objective frame. Moreover, mathematically advanced ways of obtaining the different alternatives or the so-called Pareto front are presented in this work together with a systematic approach of evaluating the different criteria.

2.3. Multi-objective optimisation

The general formulation of a multi-objective optimisation problem can be formulated as:

$$\min_{\mathbf{z}} \{J_1, J_2, \dots, J_m\} \text{ subject to: } \mathbf{z} \in S. \quad (9)$$

The feasible set S has already been defined in the previous section. The difference with a traditional dynamic optimisation problem is that more than one objective is

considered, i.e., $J_1, J_2, \dots, J_m$. In multi-objective optimisation typically no single solution exists. To determine a set of optimal solutions the concept of *Pareto optimality* is used.

Definition. A point $\mathbf{z}^* \in S$, is Pareto optimal iff there does not exist another point $\mathbf{z} \in S$, such that $J_i(\mathbf{z}) \leq J_i(\mathbf{z}^*)$ for all i and $J_j(\mathbf{z}) < J_j(\mathbf{z}^*)$ for at least one objective function j .

To obtain all the points of the Pareto set, a scalarisation approach is used in this case. The original multi-objective optimisation problem is converted in a parametric single objective optimisation problem. When this problem is solved for different values of the scalarisation parameters, a part of the Pareto front is observed. Several methods exist in literature.

- **(Convex) Weighted Sum (WS).** The Weighted Sum is the most widely used method in practice. Its formulation is:

$$\min_{\mathbf{z} \in S} J_{ws} = \sum_{i=1}^m w_i J_i \text{ with } w_i \geq 0 \text{ and } \sum_{i=1}^m w_i = 1 \quad (10)$$

with weight vector $\mathbf{w} = [w_1, \dots, w_m]^T$. There are however some drawbacks with this Weighted Sum approach (Das and Dennis, 1997). The first disadvantage is that a uniform spread of weights does not guarantee an even spread of points in the Pareto set. A second drawback is that a non-convex part of the Pareto set can not be found using the Weighted Sum approach.

- **Normal Boundary Intersection (NBI).** The NBI is developed by Das and Dennis (1998) to overcome the drawbacks of the WS. A geometrically intuitive approach is used: a plane in the objective space is built which contains all convex combinations of the individual optima, which is called the *Convex Hull of Individual Minima* (CHIM). The multi-objective problem is reformulated as maximising the distance λ of a point $\mathbf{J}_p$ on the CHIM in the direction of the utopia point along the *(quasi-)normals*. The utopia point is the point with the minimal cost values for each of the different objectives. Points closest to the utopia point on the boundary of the feasible objective space are expected to be Pareto optimal. As solutions are required to lie on the direction of the quasi-normals, extra equality constraints are introduced, while the original constraints have to be satisfied. Mathematically this is expressed as the following optimisation problem:

$$\max_{\mathbf{z} \in S, \lambda} \lambda \quad (11)$$

$$\text{subject to : } \Phi \mathbf{w} - \lambda \Phi \mathbf{e} = \mathbf{J}(\mathbf{z}) - \mathbf{J}^* \quad (12)$$

where Φ is the $m \times m$ pay-off matrix in which the i -th column is $\mathbf{J}(\mathbf{z}_i^*) - \mathbf{J}^*$, with $\mathbf{z}_i^*$ being the minimiser of the i -th objective J_i and $\mathbf{J}^*$ being the utopia point, which contains the minima of the individual objectives $J_i(\mathbf{z}_i^*)$. $\mathbf{w}$ is a vector of weights such that $\sum_{i=1}^m w_i = 1$ with $w_i \geq 0$, and $\mathbf{e}$ is a vector containing all ones. Now, $\Phi \mathbf{w}$ describes a point in the CHIM and $-\Phi \mathbf{e}$ defines the (quasi-)normal to the CHIM pointing towards the origin. When the points on the CHIM are selected with an equal spread (via a uniform distribution of $\mathbf{w}$), also an even spread on the Pareto frontier in the objective space is obtained. This method is capable of obtaining points in the non-convex part of the Pareto set. However, non-Pareto optimal solutions can also be returned.

- **(Enhanced) Normalised Normal Constraint ((E)NNC).** NNC as introduced by Messac et al. (2003), employs similar ideas as NBI, but combines

them with features of the ε -constraint method (Haimes et al., 1971). This ε -constraint method minimises a single objective function J_k , while the $m - 1$ other objective functions are added as inequality constraints $J_i \leq \varepsilon_i$, which are interpreted as hyperplanes reducing the feasible criterion space. After normalisation of the objectives, NNC also first constructs a plane through all individual minima (called here, the *utopia hyperplane*). Then NNC minimises a selected (normalised) objective $\bar{J}_k$, given the original constraints, and while additionally reducing the feasible space by adding $m - 1$ hyperplanes through a selected point $\bar{\mathbf{J}}_p$ in the utopia plane. These hyperplanes are chosen perpendicular to each of the $m - 1$ *utopia plane vectors*, which join the individual minimum $\bar{\mathbf{J}}(\mathbf{z}_k^*)$ corresponding to the selected objective $\bar{J}_k$, with all other individual minima $\bar{\mathbf{J}}(\mathbf{z}_i^*)$. Hence, this approach leads to an additional set of inequality constraints:

$$\min_{\mathbf{z} \in \mathcal{S}} \bar{J}_k \quad (13)$$

$$\text{subject to : } (\bar{\mathbf{J}}(\mathbf{z}_k^*) - \bar{\mathbf{J}}(\mathbf{z}_i^*))^T (\bar{\mathbf{J}}(\mathbf{z}) - \bar{\mathbf{J}}_p) \leq 0 \\ i = 1 \dots m, i \neq k. \quad (14)$$

As in NBI, evenly distributed points on the utopia plane $\bar{\mathbf{J}}_p$ can be selected by a uniform variation of a vector $\mathbf{w}$, which also ensures an even spread on the Pareto set. The enhanced normalised normal constraint (Sanchis et al., 2008) uses adapted normalisation schemes that in the case of more than two objectives lead to a uniform and unique representation of the Pareto set. This method is also capable of detecting points in the non-convex part. Both the NBI and (E)NNC methods sometimes return solutions which are non-Pareto optimal. These solutions can be filtered out using a filter algorithm (Messac et al., 2003; Logist et al., 2010a).

Besides scalarisation approaches there also exist a variety of evolutionary multi-objective optimisation approaches (Rangaiah, 2009). One of the most widely known algorithm and successfully applied is NSGA-II (Deb et al., 2002).

- **NSGA-II** In every iteration step i of the algorithm there is a parent population of size N , P_i , from this population a new offspring population Q_i is generated using the principles of selection, crossover and mutation. These two populations are subsequently combined in R_i . This combined population is sorted according to nondomination. Every solution is given a fitness or rank equal to its nondomination level. Solutions belonging to the first front are not dominated, those from the second front are only dominated by the first front and so on. Every element of the individual fronts is also assigned a score which is called the crowding distance. This measures how close a solution is to its neighbours. The assumption is that large average crowding distances result in better diversity in the population. The new parent population is then filled with the solutions belonging to the best nondominated set. If these number less than N , then P_{i+1} is filled with subsequent nondominated fronts in order of their ranking and crowding distance. Individuals are selected if their rank is lesser than others or if their crowding distance is larger than the others (only in case of equal rank). For a more detailed description of the algorithm, the interested reader is referred to Deb et al. (2002).

A high number of successful applications have been reported for evolutionary approaches (Rangaiah, 2009). However, they are less suited when a large number of degrees of freedom have to be optimised or other constraints than simple bounds have to be satisfied.

The software package employed to carry out the proposed strategy, is `ACADO Multi-objective` (Logist et al., 2010a), which is the multi-objective extension of the `ACADO toolkit` (www.ACADOtoolkit.org) (Houska et al., 2011). In this package the described scalarisation techniques have been implemented in a user-friendly way. For a detailed description, the interested reader is referred to the references.

2.4. Solution strategy

It is reported in literature (Franceschini and Macchietto, 2008a) that the most widely used criterion for optimal experiment design is the D-criterion. This fact does not mean that it is always the best tool for the case study at hand. The important question which remains for the experimenter is thus *which criterion to choose*. In order to allow an experimenter to make a sound decision which criterion he/she should use, the following strategy is proposed. In order to summarise the presented approach Figure 1 is depicted.

The user first has to determine which criteria he/she takes into consideration as a preliminary step to the design process (*Step 0*). Based on the recent review of Franceschini and Macchietto (2008a), all criteria described in Section 2.2 are employed in this work. These criteria are not exhaustive, if other criteria are taken into consideration, the described approach can easily be extended. In *Step 1*, the experimenter designs an optimal experiment for each criterion taken into consideration. For each of the designed experiments the scores for all other criteria are evaluated as *Step 2* and preferably tabled in a matrix format in *Step 3*. This matrix allows the experimenter to check all the expected properties for each of the designed experiments, e.g., what is the expected correlation between the criteria, what about the conditioning of the problem or how is the accuracy of the most uncertain parameter. To enhance this insight further and to visualise similarity between different criteria, a radar chart or spider plot (Miettinen, 1999) is proposed as an addition to the tabling of the results. In *Step 4* the experimenter analyses the obtained results. But, as each criterion is only a scalar objective function, certain trade-offs are expected between the proposed criteria, e.g., the ME-criterion tries to improve the conditioning of the problem but might result in large confidence regions, the D-criterion on the other hand minimises the global uncertainty volume and can thus conflict with the ME-criterion. To systematically study the trade-off and to gain insight in the process a multi-objective approach can be adopted in *Step 5*. This allows the experimenter to generate a Pareto front for a subset of selected criteria. Also these models are usually used for optimising performance, track a steady state, so to reduce the down time and minimise the influence of the experiment time, a multi-objective approach can be adopted. Sometimes several objectives are pursued without the experimenter recognising them as such. In Franceschini and Macchietto (2008b), the authors introduce several criteria for parameter accuracy and decorrelation. The approach used was to set, e.g., the information content as a constraint while minimising the correlation or vice versa. This does not allow for a systematic evaluation of the expected, intrinsic trade-offs. When no trade-offs are present one can skip *Step 5* and directly start with performing the experiments in *Step 6*. After the results of experiments are obtained, a parameter estimation procedure has to be performed. This leads to new values for the parameters. Depending on these results, e.g., the magnitude of the individual confidence bounds, several next steps are possible: restart the process with different criteria (*Step 0*), start designing experiments based on these new parameter values (iterate from *Step 1*), generate another trade-off curve for other criteria of interest (*Step 5*) or perform a new experiment to improve the current parameter estimate (*Step 6*).

3. Case studies

In this section the two case studies are introduced. The first case study is a fed-batch bioreactor. The second case study is a Lotka Volterra fishing problem.

3.1. Case 1: fed-batch bioreactor

This case study considers a well-mixed fed-batch bioreactor (Versyck and Van Impe, 1999), which is a well known benchmark problem used in the field of optimal experiment design. The dynamic model equations are given by:

$$\frac{dC_s}{dt} = -\sigma C_x + \frac{u}{V} C_{s,in} - \frac{u}{V} C_s \quad (15)$$

$$\frac{dC_x}{dt} = \mu C_x - \frac{u}{V} C_x \quad (16)$$

$$\frac{dV}{dt} = u \quad (17)$$

in which C_s [g/L] is the concentration limiting substrate, C_x [g/L] the biomass concentration and V [L] the bioreactor volume. u [L/h] is the volumetric feed rate, containing a substrate concentration $C_{s,in}$. The specific growth rate studied in this case is the monotonic *Monod* type. The mathematical expression is:

$$\mu = \mu_{max} \frac{C_s}{K_s + C_s} \quad (18)$$

The substrate consumption rate is given by the following linear law:

$$\sigma = \mu/Y_{X|S} + m \quad (19)$$

where $Y_{X|S}$ is the yield and m the maintenance factor. The parameter values are given in Table 1. The amount of biomass X and substrate S are given by: $C_x \cdot V = X$ and $C_s \cdot V = S$. The initial amount of biomass available is 10.5 g. The initial volume without substrate $V^*(0) = 7$ L. As the maximum volume $V_{max} = 10$ L and the total amount of limiting substrate available is $\alpha = 1500$ g, the initial condition of the substrate and the volume, $S(0)$ and $V(0)$ are connected as follows: $S(0) + C_{s,in}(V_{max} - V(0)) = \alpha$. The feed rate u and the end time t_f are constrained to:

$$0 \leq u \leq 1 \text{ L/h} \quad (20)$$

$$5 \leq t_f \leq 40 \text{ h} \quad (21)$$

In addition to the optimal experiment design objectives, an extra objective is considered in this case, namely maximising the biomass production expressed as $J_p = -X(t_f)$. This allows to study the possible trade-off between information content and production which occurs in an industrial environment. Unless otherwise stated a Multiple Shooting approach is used with 30 control intervals. The integrator used is a Runge-Kutta 78 with the integrator tolerance set to 10^{-8} . The resulting non-linear program is solved by an SQP method with a KKT tolerance set to 10^{-8} . To obtain the Pareto fronts 21 Pareto points are requested.

Table 1: Parameter values for the fed-batch bioreactor.

μ_{max}	0.1 h ⁻¹	K_s	1 g/L
m	0.29 g/g	$Y_{X S}$	0.47 g/g
$\sigma_{C_s}^2$	10 ⁻² g ² /L ²	$\sigma_{C_x}^2$	6.25 × 10 ⁻⁴ g ² /L ²

3.2. Case 2: Lotka Volterra fishing problem

This case study considers the Lotka Volterra fishing problem (Sager et al., 2009; Sager, 2012). In this model a fishing strategy is determined in a fixed time horizon to bring both the biomass of the prey and the predator to a certain predefined optimal steady state. In the literature this model is used to study mixed integer optimal control problems as the control input can be considered to be binary. In more recent work this model is employed to study sampling decisions in optimal experiment design (Sager, 2012). The model equations are:

$$\frac{dx_1}{dt} = x_1 - p_1 x_1 x_2 - 0.4 x_1 u \quad (22)$$

$$\frac{dx_2}{dt} = -x_2 + p_2 x_1 x_2 - 0.2 x_2 u \quad (23)$$

where x_1 is the biomass of the prey and x_2 the biomass of the predator. u is the fishing control, which is assumed to be binary: $u \in \{0, 1\}$. The initial conditions are known to be: $x_1(0) = 0.5$ and $x_2(0) = 0.7$, furthermore the end time is fixed at $t_f = 12$. The nominal values of parameters are $p_1 = 1$ and $p_2 = 1$. In Sager et al. (2009), the following tracking objective function is employed:

$$\min \int_0^{t_f} (x_1 - 1)^2 + (x_2 - 1)^2 dt \quad (24)$$

In order to solve the optimisation problems, a Multiple Shooting approach is adopted. The number of used control intervals is 60. This binary control problem is solved in a relaxed version, $u \in [0, 1]$ is solved. Afterwards the sum-up rounding strategy (Sager et al., 2009) is applied to end up with a binary control formulation. In order to extend the approach to multi-objective optimisation, the strategy by Logist et al. (2010b) is adopted. The used integrator is a BDF with tolerance set to 10^{-6} . The non-linear program is solved by an SQP routine with tolerances set to 10^{-5} .

4. Results and discussion

For both case studies the criteria described in Section 2.2 are used (Step 0). Some of the proposed experiments are studied further on a simulation level. Gaussian, white noise is sampled from a normal distribution and added to the simulated measurements. The variances are taken equal to the measurement output variances. These simulated measurements are then used in a parameter estimation procedure. To solve the parameter procedure the same numerical methods and conditions are used as in the optimisation of the different experiment designs.

4.1. Case study 1: fed-batch bioreactor

In Table 2 the results of Steps 1 and 2 are depicted. This pay-off table provides the values for each of the criteria horizontally for each optimal experiment which are listed vertically (Step 3). Listing these values in such a way allows the practitioner to study the differences between the criteria in a systematic way.

The analysis of the results (Step 4) show that three different groups of criteria can be distinguished. First, there are the A-, D-, M-criterion and the production which are quite similar. This can also be seen in Figure 2, where the relative scores for the different criteria are depicted in a spider plot. The A-, D-, and M-criterion score relatively high on their respective criteria. However, they all result in experiments where the parameters are highly correlated and the system is badly conditioned as can be seen in the values for the ME-criterion and the correlation coefficient.

Second, there are the criteria which do not target the parameter accuracy but the conditioning of the problem or the correlation. Both the ME-criterion and the correlation experiments result in significantly lower values for production and the parameter accuracy criteria as indicated by the A-, D- and M-criterion. Note also that the optimal experiment for the ME-criterion has quite a high correlation and the correlation experiment has a relatively large condition number.

Apart from these two groups there is the E-criterion. This experiment has a low value for the A-criterion, however, its D-criterion value is remarkably high. When the ME-criterion and the correlation are taken into account a difference with the other parameter accuracy criteria is observed. The E-criterion has the second best result for both criteria. This observation is also clearly illustrated in Figure 2 which distinguishes the E-criterion from the other parameter accuracy criteria as it does not show large trade-offs. These observations are summarised in a comparative graph in Figure 2. A multi-objective approach is now used to mitigate the drawbacks described in the previous section and to study the trade-offs between some criteria evident in Table 2. This evaluation is Step 5 of the described approach. Based on the resulting Pareto fronts, experiments are subsequently chosen which are examined on a simulation level (Steps 6 and 7).

Table 2: Pay-off matrix for the different criteria for each of the designed experiments for the fed-batch bioreactor case, each row gives the scores of the designed experiment for all criteria taken into consideration.

Score for: Experiment:	Production	A-criterion	D-criterion	E-criterion	ME-criterion	M-criterion	Correlation
Production	3.02 10²	1.28 10 ¹⁰	3.71 10 ¹²	2.90 10 ²	4.40 10 ⁷	1.10 10 ⁴	9.87 10 ⁻¹
A-criterion	3.01 10 ²	1.32 10¹⁰	6.23 10 ¹²	4.72 10 ²	2.80 10 ⁷	2.59 10 ⁴	9.91 10 ⁻¹
D-criterion	2.96 10 ²	9.49 10 ⁹	3.60 10¹⁴	3.79 10 ⁴	2.50 10 ⁵	1.16 10 ⁶	9.84 10 ⁻¹
E-criterion	2.78 10 ²	2.50 10 ⁸	1.10 10 ¹⁴	4.40 10⁵	5.66 10 ²	4.66 10 ⁵	2.37 10 ⁻¹
ME-criterion	1.54 10 ¹	1.65 10 ⁴	7.33 10 ⁵	4.45 10 ¹	3.69 10²	7.23 10 ²	6.19 10 ⁻¹
M-criterion	2.94 10 ²	9.06 10 ⁹	3.04 10 ¹⁴	3.36 10 ⁴	2.70 10 ⁵	1.27 10⁶	9.87 10 ⁻¹
Correlation	1.52 10 ¹	3.27 10 ⁴	3.32 10 ⁵	1.02 10 ¹	3.21 10 ³	1.02 10 ¹	8.36 10⁻²

4.1.1. Combining the ME-criterion and maximal production

A first combination of criteria includes production and the modified E-criterion. In Figure 3 the resulting Pareto set is shown for $J_1 = \frac{\lambda_{max}(\mathbf{F})}{\lambda_{min}(\mathbf{F})}$ and $J_2 = -X(t_f)$. Note that every optimisation is expressed as a minimisation problem. A trade-off is clearly observed between on the one hand maximising production, which results in a large, undesired condition number and on the other hand minimising the condition number, resulting in a sharp decrease of production. The NNC is however, able to represent the Pareto set more accurately than the WS. The latter method does not represent the production decreasing part of the Pareto front properly.

The resulting states of the five weight combinations are depicted in Figure 4. The shift in focus of the different objectives is clearly observable in the states. The optimal production control means supplying the maximal amount of substrate at the start and continuing in a batch mode. When the focus is gradually shifted to minimising the condition number, the amount of initial substrate decreases until the case of the minimal condition number where a minimal amount of substrate is available at the start. Figure 4 also illustrates, that each of the combinations has a different end time.

The resulting parameter values of the simulation study can be seen in Table 3. Three different experiments are simulated: the two individual optima and the intermediate case with $w = [0.5 \ 0.5]^T$ combination. It can be seen that the smallest confidence bounds are found in the production case. However as depicted in Figure 5, this leads to a very ill-conditioned parameter estimation problem. To overcome this ill-conditioning the ME-criterion can be used but as the result in Table 3 indicates, this leads to a larger confidence interval. This can be explained by how the optimisation of the ME-criterion works. This criterion tries to obtain diagonal elements in the Fisher information that are equal. The non-diagonal elements are being forced to 0. This however does not guarantee an experiment with small confidence bounds. It only allows to find a better conditioned parameter estimation problem. This issue is illustrated in Figure 5. Both the extremal cases of weight choices are shown together with the intermediate case given by $w = [0.5 \ 0.5]^T$. Figure 5 depicts the contours of the corresponding cost surfaces. As can be seen, when the focus is on production a bad conditioning is present based on the elongated contours, which do not even close in the current plot. Alternatively, when the ME-criterion is focussed on, the contours become more and more circular and close, reflecting the better conditioning. In case of the trade-off experiment an ellipsoidal joint confidence interval is obtained whereas the individual confidence intervals are in the same order of magnitude as the production case. The effect of the better conditioning is also present in Table 3 for $w = [0.5 \ 0.5]^T$, which leads to similarly equal accurate parameter results as in the production case, but with a better defined optimum. To improve the accuracy even further dedicated parameter accuracy criteria are considered in the following combination of criteria.

Table 3: Parameter values for simulated experiments for the optimal ME-criterion ($w = [1 \ 0]^T$), the maximal production case ($w = [0 \ 1]^T$) and a trade-off case ($w = [0.5 \ 0.5]^T$).

Operating condition	μ_{max}	K_s
$w = [1 \ 0]^T$	$9.218 \ 10^{-2} \pm 2.267 \ 10^{-2}$	$1.315 \ 10^0 \pm 5.634 \ 10^{-1}$
$w = [0.5 \ 0.5]^T$	$9.988 \ 10^{-2} \pm 4.253 \ 10^{-4}$	$9.976 \ 10^{-1} \pm 3.390 \ 10^{-2}$
$w = [0 \ 1]^T$	$1.000 \ 10^{-1} \pm 1.090 \ 10^{-4}$	$1.016 \ 10^0 \pm 9.056 \ 10^{-2}$

4.1.2. Combining the ME-criterion and the D-criterion

Combining the Modified E-criterion, $J_1 = \frac{\lambda_{max}(\mathbf{F})}{\lambda_{min}(\mathbf{F})}$, and the D-criterion, $J_2 = -\det(\mathbf{F})$ is an interesting case. The combination allows obtaining more circular joint confidence intervals while also minimising the global volume. The Pareto set is depicted in Figure 6. A clear trade-off is observed between the two criteria. When shifting the focus from maximising the determinant to minimising the condition number, at first the condition number can be decreased while loosing one order of magnitude in the determinant. At a certain point however, when the focus is more shifted to minimising the condition number a sharp decrease in the determinant is found. When comparing the ENNC and the WS one notices that the ENNC is capable of representing the Pareto front more accurately than the WS. This clearly illustrates that the WS is not able to return points in the non-convex part of the Pareto front.

The resulting states for the five different weight combinations are displayed in Figure 7. The feeding profile for all weight combination except maximising the determinant exhibit a similar behaviour. Initially, there is a small amount of feeding to the reactor, which is also visible in the evolution of the volume. At some time point however there is a increase to the maximal feed rate. This behaviour can also be noticed in the substrate concentration. From the amount of biomass formed at

the end, it can be concluded that there is only a slight decrease in production when optimising the determinant.

To validate the results simulation experiments are performed. The resulting values for three different cases are shown in Table 4. As in Table 3, using only the ME-criterion does not necessarily lead to accurate estimations of the individual parameters. This is illustrated by the broad confidence intervals. The D-criterion as seen in Figure 8, also suffers from ill-conditioning and highly correlated parameters. However, the accuracy in which K_s can be estimated is better than in the previous case, as the D-criterion tries to have a globally accurate design. When using the trade-off experiment a slightly worse result is found for μ_{max} but a better result for K_s . This illustrates that using a multi-objective approach allows the user to come up with trade-off experiments which allow to smoothen the drawbacks of both criteria. The better conditioning and the global accuracy criterion allows the user in this case to find a very accurate estimate for K_s . Figure 8 depicts the contours of the corresponding cost surfaces. As can be seen, when the D-criterion is focussed on an ill-conditioned and highly correlated cost surface is found. Alternatively, when better conditioning is focussed on, a much larger joint confidence region is observed. Remark however that the correlation is reduced as compared to the D-criterion. In this case, it is also illustrated that the Fisher Information Matrix is only an approximation of the parameter uncertainty. It can be seen that the individual confidence intervals underestimate the joint confidence interval. When the ME-criterion is focussed on the same result as Figure 5 is obtained. This illustrates again the potential drawback of the ME-criterion, i.e., results with more circular but bigger joint confidence intervals.

Table 4: Parameter values for simulated experiments for the optimal ME-criterion ($w = [1\ 0]^T$), the optimal D-criterion case ($w = [0\ 1]^T$) and a trade-off case ($w = [0.5\ 0.5]^T$).

Operating condition	μ_{max}	K_s
$w = [1\ 0]^T$	$9.382\ 10^{-2} \pm 2.191\ 10^{-2}$	$1.069\ 10^0 \pm 4.664\ 10^{-1}$
$w = [0.5\ 0.5]^T$	$1.008\ 10^{-1} \pm 4.153\ 10^{-4}$	$1.0025\ 10^0 \pm 8.576\ 10^{-3}$
$w = [0\ 1]^T$	$1.001\ 10^{-1} \pm 1.552\ 10^{-4}$	$1.0088\ 10^0 \pm 1.4932\ 10^{-2}$

4.1.3. Combining the M-criterion and the decorrelation criterion

In this section the decorrelation criterion is used in combination with the M-criterion, $J_1 = -\min F_{ii}$. As this is a two parameter problem, there is only one correlation coefficient. Expressed in terms of the elements of the Fisher Information Matrix this criterion is $J_2 = C_{12} = \frac{-F_{12}}{\sqrt{F_{11}}\sqrt{F_{22}}}$. In contrast with the approach used in Franceschini and Macchietto (2008b), both the decorrelation and parameter accuracy are considered as objectives. An advantage of this approach is that the entire resulting Pareto front is displayed in Figure 9. It can be seen that the NBI method generates points which are not Pareto optimal and which are detected by the Pareto filter and subsequently removed. Even after filtering enough Pareto points remain to illustrate the trade-off between accuracy and decorrelation. Also note the fact that there is a discontinuity in the Pareto front.

The resulting controls and states are depicted in Figure 10. Interesting to note is that the input profile of the decorrelation criterion resembles the input profile for the ME-criterion, although a longer end time is found in the decorrelation case.

The resulting parameter values, after a simulation experiment, are shown in Table 5. The results indicate clearly the existence of the trade-off that was already visible in the Pareto front. When trying to maximise the information content the smallest confidence bounds are found, however if the decorrelation is reduced from

almost 1 to around 0.30, larger confidence bounds are found but still within a same order of magnitude. When driving the correlation to the smallest value possible, even broader confidence regions are found. Note also that this evolution is very similar to the one found in the case of the combination of the ME-criterion and the D-criterion. To evaluate the effect of decorrelation, the joint confidence bounds are also displayed in Figure 11. It can be seen that the more these are driven to their optimal value, the more the axes of the ellipse are aligned with the axes of the parameter values. The price to pay, however, is a large decrease in information content. The joint confidence region is notably larger in the least correlation case.

Table 5: Parameter values for simulated experiments for the M-criterion ($w = [10]^T$), the minimal correlation case ($w = [0\ 1]^T$) and a trade-off case ($w = [0.5\ 0.5]^T$).

Operating condition	μ_{max}	K_s
$w = [10]^T$	$9.999\ 10^{-2} \pm 2.084\ 10^{-4}$	$9.993\ 10^{-1} \pm 1.745\ 10^{-2}$
$w = [0.5\ 0.5]^T$	$9.991\ 10^{-2} \pm 4.137\ 10^{-4}$	$9.939\ 10^{-1} \pm 3.149\ 10^{-2}$
$w = [0\ 1]^T$	$1.057\ 10^{-1} \pm 1.139\ 10^{-2}$	$1.060\ 10^0 \pm 1.220\ 10^{-1}$

4.2. Case study 2: Lotka Volterra fishing problem

In Table 6 the results of Step 1 and 2 are summarised (Step 3). An analysis of these results (Step 4) reveals that four different types of results can be distinguished. The same results are also represented as a radar chart in Figure 2. The A- and D-criterion can be grouped together as they show very similar results on all criteria with the exception of the correlation value. The A-criterion has the highest correlation value whereas the D-criterion only scores 0.146. For the other criteria both the A- and D-criterion experiments score almost identically. This is not too surprising as both try to reduce the global uncertainty on the parameters.

On the other hand, there is the group of the E-, ME- and M-criterion. In their respective criteria they form the top three. When the D- and A-criterion are taken into consideration, it is observed that the E-, ME- and M-criterion optimal experiments score similarly. Note that compared to the A- and D-criterion optimal experiments, these experiments score lower. This observation can be explained from the fact that these criteria try to estimate the most uncertain parameter (E- and M-criterion) or try to distribute the fault evenly between the parameters (ME-criterion) and, thus, do not focus on the uncertainty of the other parameters.

The correlation criterion stands somewhat apart from the others as it scores the worst as optimal experiment on the other design criteria except for the ME-criterion. This means that trying to reduce the correlation between the parameters will result in an experiment which will have a larger uncertainty on the parameter values.

To conclude the overview there is the binary benchmark tracking criterion. It can easily be seen that this results in very poor scores for all the different OED-criteria with the exception of the ME-criterion score. Note that no experiment results in a very ill-conditioned experiment. All the scores for the ME-criterion are in the range of 1.00 and 4.90 with the A- and D-criterion resulting in a slightly worse condition number. This indicates that this case study is a better conditioned problem compared to the previous one. As there is not a single optimal experiment that really stands out and to mitigate the drawbacks described in the previous section and to study the trade-offs between some criteria evident in Table 6, the multi-objective approach is employed (Step 5). The used measurement error variances, $\sigma_{x_1}^2$ and $\sigma_{x_2}^2$

is 0.09 for both states. Based on the obtained Pareto sets, experiments are chosen which are examined in silico (Step 6 and 7).

Table 6: Pay-off matrix for the different criteria for each of the designed experiments for the Lotka Volterra fishing case, each row gives the scores of the designed experiment for all criteria taken into consideration.

Score for: Experiment:	Tracking	A-criterion	D-criterion	E-criterion	ME-criterion	M-criterion	Correlation
Tracking	1.34 10^0	5.60 10^1	7.23 10^2	2.02 10^1	1.76 10^0	2.34 10^1	2.27 10^{-1}
A-criterion	2.05 10^1	1.62 10^2	3.68 10^3	2.74 10^1	4.90 10^0	3.33 10^1	3.74 10^{-1}
D-criterion	2.24 10^1	1.52 10^2	4.02 10^3	3.41 10^1	3.45 10^0	3.52 10^1	1.46 10^{-1}
E-criterion	1.83 10^1	9.77 10^1	2.38 10^3	4.64 10^1	1.11 10^0	4.74 10^1	-4.20 10^{-2}
ME-criterion	1.85 10^1	9.59 10^1	2.30 10^3	4.63 10^1	1.00 10^0	4.68 10^1	-2.61 10^{-2}
M-criterion	1.95 10^1	1.04 10^2	2.59 10^3	4.14 10^1	1.52 10^0	5.29 10^1	2.04 10^{-1}
Correlation	5.10 10^0	6.47 10^1	9.83 10^2	2.43 10^1	1.66 10^0	2.43 10^1	-1.46 10^{-6}

4.2.1. Combining the tracking objective and the D-criterion

From a practical point of view, combining an existing, desired, tracking objective, $J_1 = \int_0^{t_f} (x_1 - 1)^2 + (x_2 - 1)^2 dt$ with a criterion which allows to further refine the existing model is an evident choice. The D-criterion, $J_2 = -\det(\mathbf{F})$ is chosen as it allows the enhancement of the global accuracy of the parameter estimate. Performing the D-criterion experiment results however in the worst tracking result available as evident in Table 6. The resulting Pareto front, obtained with the NBI, is displayed in Figure 12.

To illustrate the trade-off between the tracking objective and the D-criterion the resulting states and control inputs are depicted in Figure 13 for three different points of the Pareto front. The corresponding states and control input for three points are depicted in Figure 13. When the tracking objective is focussed on a clear steady state behaviour is reached after four time units, when the focus is divided between the tracking and the D-criterion an oscillation takes place in both states. If the D-criterion is optimised the resulting states oscillate even more than in the trade-off case. Note that the control inputs in Figure 13 are the result of the sum-up rounding strategy.

By performing the proposed experiments in a simulation study the following parameter estimates are obtained (see Table 7). From this table one can see that the tracking criterion results in large uncertainty intervals. When the objective is shifted to the D-criterion these intervals become smaller and are smallest in the case when only the D-criterion is focussed on. The joint confidence intervals can be seen in Figure 14. As expected from the individual confidence intervals, the tracking criterion results in the largest joint confidence region. The trade-off case and the D-criterion result however in almost equally large joint confidence regions. These regions are clearly smaller than the tracking case. An interesting issue arises in the trade-off case. It results in an experiment with a higher correlation value, i.e., 0.47, than the two individual optimal cases. In Figure 14, this effect can be observed by the fact that the joint confidence region is more tilted and not aligned with coordinate axes. To take this correlation into account, a natural extension is to add it as an extra objective.

4.2.2. Combining the tracking objective, correlation and the D-criterion

As Steps 6 and 7 resulted in a too correlated experiment an additional iteration, incorporating the correlation as an extra objective, is performed. The objectives

Table 7: Parameter values for simulated experiments for the minimal tracking value ($w = [1\ 0]^T$), the maximal D-criterion case ($w = [0\ 1]^T$) and a trade-off case ($w = [0.5\ 0.5]^T$).

Operating condition	p_1	p_2
$w = [1\ 0]^T$	$9.978\ 10^{-1} \pm 3.778\ 10^{-2}$	$9.973\ 10^{-1} \pm 3.169\ 10^{-2}$
$w = [0.5\ 0.5]^T$	$1.001\ 10^0 \pm 2.666\ 10^{-2}$	$1.001\ 10^0 \pm 1.956\ 10^{-2}$
$w = [0\ 1]^T$	$9.823\ 10^{-1} \pm 2.552\ 10^{-2}$	$1.009\ 10^0 \pm 1.450\ 10^{-2}$

are $J_1 = \int_0^{t_f} (x_1 - 1)^2 + (x_2 - 1)^2 dt$, $J_2 = -\frac{F_{12}}{\sqrt{F_{11}}\sqrt{F_{22}}}$ and $J_3 = -\det(\mathbf{F})$.

In Figure 12, the three objective Pareto front can be seen. In comparison with the previous Pareto front two additional points are of interest, i.e., the case when only correlation is focussed on, $[0\ 1\ 0]^T$ and the trade-off point $[0.33\ 0.33\ 0.33]^T$. Figure 13 illustrates the resulting states and the control input. From Figure 13, it is clear that minimising the correlation results in an experiment that has the second lowest value for the tracking objective. This is also clearly evident in Table 6. The optimal correlation case also results in a control action which is clearly highly nervous. When the trade-off point $w = [0.33\ 0.33\ 0.33]^T$ is taken, the tracking result is clearly less accurate but it outperforms the trade-off case of the previous section. The control action is also quite similar to the two dimensional $w = [0.5\ 0.5]^T$ case. In Table 8 the results of the simulation experiments are given. From these results, it is clear that only focussing on decorrelation clearly results in broad individual confidence intervals. In comparison with the result from Table 7, one can conclude that this results in confidence intervals that are slightly better than the tracking criterion. Based on the results of Table 6, this observation is to be expected as the correlation criterion scores poorly on all other criteria. The trade-off case, $w = [0.33\ 0.33\ 0.33]^T$ results in individual confidence bounds larger than in the previous case. The resulting joint confidence regions are also illustrated in Figure 14. Minimising the correlation results in a large confidence region, only the tracking criterion results in a larger joint confidence region. As also visible in the individual confidence intervals, the trade-off case results in a joint confidence region greater than the D-criterion and the trade-off case of the previous section. The correlation, indicated by the alignment of the joint confidence region, is in this trade-off result however lower than previous.

Table 8: Parameter values for simulated experiments for the minimal correlation ($w = [0\ 1\ 0]^T$) and the trade-off case ($w = [0.33\ 0.33\ 0.33]^T$).

Operating condition	p_1	p_2
$w = [0\ 1\ 0]^T$	$9.844\ 10^{-1} \pm 3.247\ 10^{-2}$	$1.023\ 10^0 \pm 2.703\ 10^{-2}$
$w = [0.33\ 0.33\ 0.33]^T$	$1.022\ 10^0 \pm 2.918\ 10^{-2}$	$1.018\ 10^0 \pm 2.213\ 10^{-2}$

4.2.3. Combining the A-criterion and the E-criterion

The A-criterion focusses on the global accuracy of the parameters, but can suffer when encountering correlated parameters (Franceschini and Macchietto, 2008a). Based on the results in Table 6, one sees that with respect to the most uncertain parameter (E- and M-criterion), the A-criterion scores poorly. Both the E- and M-criterion score similarly with respect to the global accuracy criteria, but only the E-criterion seems also to result in an almost uncorrelated experiment. So, to overcome the correlation problem of the A-criterion, the E-criterion is used. The two used criteria are the A-criterion, $J_1 = -\text{trace}(\mathbf{F})$ and the E-criterion $J_2 = -\lambda_{\min}(\mathbf{F})$. The resulting Pareto front is depicted in Figure 15.

The resulting states and control action are illustrated in Figure 16. All experiments have in common that they score very badly on the tracking criterion. All experiments also result in a highly oscillatory states. Although from first view, the experiments seem similar, different results are obtained in the parameter estimation simulation study. The resulting values are portrayed in Table 9. When the A-criterion is focussed on a parameter accuracy is obtained which is slightly worse than the result for the D-criterion in Table 7. By introducing the E-criterion as an additional objective the uncertainty regarding the most uncertain parameter is reduced while only reducing slightly the accuracy of the most certain parameter. When only the E-criterion is used, this results in an experiment with almost equal uncertainty. The drawback however is that the confidence interval for the most certain parameter increased.

One can find the joint confidence intervals in Figure 16. The E-criterion results in a quite large almost circular joint confidence region. When the focus is put more on the global accuracy, a joint confidence region is obtained that is completely encompassed by the one for the E-criterion. The joint confidence region of the A-criterion is more tilted than the two others, illustrating the higher correlation value. The area covered by the A-criterion and the trade-off confidence region appears to be similar, with the difference however of having less correlation and a better conditioning. This illustrates again the strength of combining several optimal experiment design and study them systematically in a multi-objective way.

Table 9: Parameter values for simulated experiments for the maximal A-criterion, ($w = [1\ 0]^T$), the maximal E-criterion, ($w = [0\ 1]^T$) and the trade-off case ($w = [0.5\ 0.5]^T$).

Operating condition	p_1	p_2
$w = [1\ 0]^T$	$1.013\ 10^0 \pm 2.991\ 10^{-2}$	$9.924\ 10^{-1} \pm 1.519\ 10^{-2}$
$w = [0.5\ 0.5]^T$	$9.791\ 10^{-1} \pm 2.395\ 10^{-2}$	$1.006\ 10^0 \pm 1.561\ 10^{-2}$
$w = [0\ 1]^T$	$9.789\ 10^{-1} \pm 2.746\ 10^{-2}$	$9.789\ 10^{-1} \pm 2.709\ 10^{-2}$

4.3. Using more than three objectives

In the previous case studies two or at most three objectives have been considered. The main reason for employing two to three objectives is that it remains easily presentable. In this section, it is illustrated that the presented approach is also feasible for more objectives.

In Figure 18 the Pareto front of four objectives for both cases is presented. In both cases a front of ten points is computed. In the case of more than three objectives it is proposed to calculate first a coarse approximation of the Pareto front. This is motivated by the fact that the number of Pareto points and number of objectives scale in a factorial way. Therefore, a first coarse approximation is calculated, depending on the preference of the experimenter, a more precise calculation of the Pareto front can be executed in a second phase. In order to aid the experimenter in his decision-making an interactive multi-objective method can be used, e.g. the NIMBUS method (Miettinen, 1999).

For the first case study the following criteria are used: correlation, M-, D- and ME-criterion. In the upper part of Figure 18 the resulting Pareto front is displayed. Note again the sharp trade-off between on the one hand the M- and D-criterion and on the other hand the correlation and ME-criterion. However, the computation of the Pareto front reveals the existence of trade-off solution as experiment 5 where all scores are within the 50-80 % range. In the second case study the following criteria are employed: tracking, A-, E- and D-criterion. The resulting Pareto front

is shown in the lower part of Figure 18. From the Pareto front one can conclude that criteria from optimal experiment design contrast strongly with the tracking in a steady state. In addition it is illustrated that the A- and D-criterion are usually closely related, except for experiment 8.

The bar charts in Figure 18 illustrate the presented method is also feasible for more than three objectives. The main advantage however becomes representing the obtained results. If additional objectives and experiments are to be added the overview and insight for the experimenter might get lost.

5. Conclusion

The question addressed in this paper is *which criterion to choose* in an optimal experiment design. To tackle this question, a generic approach to evaluate the different criteria of interest is presented. Optimal experiment design leads to dynamic optimisation problem. These dynamic optimisation problems are solved by direct optimal control algorithms. As each criterion is only a scalar function of the Fisher Information Matrix, they all have their strengths and weaknesses. In order to mitigate these drawbacks and to study the trade-off between several criteria a multi-objective approach has been adopted. In this multi-objective approach novel scalarisation methods are employed. Furthermore, it is illustrated that the multi-objective character is sometimes ignored. In order to avoid determining a priori levels of information content and to systematically evaluate all trade-off solution, a multi-objective optimisation is proposed. This allows the decision maker to make an informed decision. To illustrate the developed approach, two case studies are presented. The first is a fed-batch bioreactor which is well known example in the field of optimal experiment design whereas the second case study is used by Sager in Sager (2012).

ACKNOWLEDGMENTS

Work supported in part by Projects, OT/10/035, PFV/10/002 OPTEC (Centre-of-Excellence Optimisation in Engineering) and KP/09/005 SCORES4CHEM knowledge platform of the Katholieke Universiteit Leuven, by the Belgian Program on Interuniversity Poles of Attraction, initiated by the Belgian Federal Science Policy Office and by industry project ACCM. Dries Telen and Ignace Tack have a Ph.D grant of the institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). Jan Van Impe holds the chair Safety Engineering sponsored by the Belgian chemistry and life sciences federation essenscia.

References

- Alberton, A., Schwaab, M., Biscaia jr., E., J.C., P., 2010. Sequential experimental design based on multiobjective optimization procedures. Chemical Engineering Science 65, 5482–5494.
- Balsa-Canto, E., Alonso, A., Banga, J., 2008. Computing optimal dynamic experiments for model calibration in predictive microbiology. Journal of Food Process Engineering 32 (2), 186–206.
- Balsa-Canto, E., Banga, J., Alonso, A., Vassiliadis, V., 2001. Dynamic optimization of chemical and biochemical processes using restricted second-order information. Computers & Chemical Engineering 25 (4-6), 539–546.

- Bauer, I., Bock, G., Korkel, S., Schlöder, J., 2000. Numerical methods for optimum experimental design in dae systems. *Journal of Computational and Applied Mathematics* 120, 1–25.
- Biegler, L., 1984. Solution of dynamic optimization problems by successive quadratic programming and orthogonal collocation. *Computers and Chemical Engineering* 8, 243–248.
- Biegler, L., 2007. An overview of simultaneous strategies for dynamic optimization. *Chemical Engineering and Processing: Process Intensification* 46, 1043–1053.
- Bock, H., Plitt, K., 1984. A multiple shooting algorithm for direct solution of optimal control problems. In: *Proceedings of the 9th IFAC world congress, Budapest*. Pergamon Press.
- Das, I., Dennis, J., 1997. A closer look at drawbacks of minimizing weighted sums of objectives for Pareto set generation in multicriteria optimization problems. *Structural Optimization* 14, 63–69.
- Das, I., Dennis, J., 1998. Normal-Boundary Intersection: A new method for generating the Pareto surface in nonlinear multicriteria optimization problems. *SIAM Journal on Optimization* 8, 631–657.
- Deb, K., Pratap, A., Agarwal, S., Meyarivan, T., 2002. A fast and elitist multi-objective genetic algorithm: NSGA-II. *IEEE Transaction on Evolutionary Computation* 6, 181–197.
- Donckels, B., De Pauw, D., Vanrolleghem, P., De Baets, B., 2010. An ideal point method for the design of compromise experiments to simultaneously estimate the parameters of several rival models. *Chemical Engineering Science* 65, 1705–1719.
- Fisher, R., 1935. *The design of experiments*. Oliver & Boyd.
- Franceschini, G., Macchietto, S., 2008a. Model-based design of experiments for parameter precision: State of the art. *Chemical Engineering Science* 63, 4846–4872.
- Franceschini, G., Macchietto, S., 2008b. Novel anticorrelation criteria for model-based experiment design: theory and formulations. *A.I.Ch.E Journal* 54, 1009–1024.
- Haimes, Y., Lasdon, L., Wismer, D., 1971. On a bicriterion formulation of the problems of integrated system identification and system optimization. *IEEE Transactions on Systems, Man, and Cybernetics SMC-1*, 296–297.
- Houska, B., Ferreanu, H., Diehl, M., 2011. ACADO Toolkit - an open-source framework for automatic control and dynamic optimization. *Optimal Control Applications and Methods* 32, 298–312.
- Körkel, S., Kostina, E., Bock, H., Schlöder, J., 2004. Numerical methods for optimal control problems in design of robust optimal experiments for nonlinear dynamic processes. *Optimization Methods and Software Journal* 19 (3-4), 327–338.
- Kud, A., Körkel, S., Maixner, S., 2010. A cubic equation of state based on saturated vapour modeling and the application of model-based design of experiments for its validation. *Chemical Engineering Science* 65, 4194–4207.
- Leineweber, D., Bauer, I., Bock, H., Schlöder, J., 2003. An efficient multiple shooting based reduced SQP strategy for large-scale dynamic process optimization. Part I: theoretical aspects. *Computers and Chemical Engineering* 27, 157–166.

- Ljung, L., 1999. System Identification: Theory for the User. Prentice Hall.
- Logist, F., Houska, B., Diehl, M., Van Impe, J., 2010a. Fast pareto set generation for nonlinear optimal control problems with multiple objectives. Structural and Multidisciplinary Optimization 42, 591–603.
- Logist, F., Houska, B., Diehl, M., Van Impe, J., 2011a. Robust multi-objective optimal control of uncertain (bio)chemical processes. Chemical Engineering Science 66, 4670–4682.
- Logist, F., Sager, S., Kirches, C., van Impe, J., 2010b. Efficient multiple objective optimal control of dynamic systems with integer controls. Journal of Process Control 20 (7), 810–822.
- Logist, F., Telen, D., Van Derlinden, E., Van Impe, J., 2011b. Multi-objective optimisation approach to optimal experiment design using acado toolkit. In: Proceedings of the 21st European symposium on computer aided process engineering. Elsevier, pp. 462–466.
- Logist, F., Van Impe, J., 2012. Novel insights for multi-objective optimisation in engineering using normal boundary intersection and (enhanced) normalised normal constraint. Structural and Multidisciplinary Optimization 45, 417–431.
- Messac, A., Ismail-Yahaya, A., Mattson, C., 2003. The normalized normal constraint method for generating the Pareto frontier. Structural and Multidisciplinary Optimization 25, 86–98.
- Miettinen, K., 1999. Nonlinear multiobjective optimization. Kluwer Academic Publishers, Boston.
- Pukelsheim, F., 1993. Optimal design of Experiments. John Wiley & Sons, Inc., New York.
- Rangaiah, G. P., 2009. Multi-objective optimization - Techniques and Applications in Chemical Engineering. World Scientific Publishing, Singapore.
- Sager, S., 2012. Sampling decisions in optimum experimental design in the light of Pontryagin's maximum principle. SIAM Journal on Control and Optimization(submitted).
URL http://www.optimization-online.org/DB_HTML/2011/05/3037.html
- Sager, S., Bock, H., Diehl, M., G., R., Schlöder, J., 2009. Numerical methods for optimal control with binary control functions applied to a Lotka-Volterra type fishing problem. In: Seeger, A. (Ed.), Recent Advances in Optimization. Vol. 563 of Lectures Notes in Economics and Mathematical Systems. Springer, Heidelberg, pp. 269–289.
- Sanchis, J., Martinez, M., Blasco, X., Salcedo, J., 2008. A new perspective on multi-objective optimization by enhanced normalized normal constraint method. Structural and Multidisciplinary Optimization 36, 537–546.
- Sargent, R., Sullivan, G., 1978. The development of an efficient optimal control package. In: Stoer, J. (Ed.), Proceedings of the 8th IFIP Conference on Optimization Techniques. Springer-Verlag, Heidelberg, pp. 158–168.
- Van Derlinden, E., Bernaerts, K., Van Impe, J., 2010. Simultaneous versus sequential optimal experiment design for the identification of multi-parameter microbial growth kinetics as a function of temperature. Journal of Theoretical Biology 264, 347–355.

Vassiliadis, V., Sargent, R., Pantelides, C., 1994a. Solution of a class of multistage dynamic optimization problems. 1. Problems without path constraints. Industrial and Engineering Chemistry Research 33, 2111–2122.

Vassiliadis, V., Sargent, R., Pantelides, C., 1994b. Solution of a class of multistage dynamic optimization problems. 2. Problems with path constraints. Industrial and Engineering Chemistry Research 33, 2123–2133.

Versyck, K., Van Impe, J., September 1998. Trade-offs in design of fed-batch experiments for optimal estimation of biokinetic parameters. In: Proceedings of the 1998 IEEE International Conference on Control Applications. Trieste, Italy.

Versyck, K., Van Impe, J., 1999. Feed rate optimization for bed-batch bioreactors: from optimal process performance to optimal parameter estimation. Chemical Engineering Communications 172, 107–124.

Walter, E., Pronzato, L., 1997. Identification of Parametric Models from Experimental Data. Springer, Paris.

Figure 1: Schematic representation of the proposed procedure for choosing an optimal experiment design criterion.

Figure 2: Relative comparison of the different optimal experiments with respect to the different criteria for the fed-batch bioreactor case (upper) and the Lotka Volterra case (lower) study.

Figure 3: Pareto front obtained by the ENNC (left) and the WS (right) of the combination ME-criterion and production.

Figure 4: Evolution substrate and biomass concentration (upper left and right), volume and controls (lower left and right) for 5 weight combinations ME-criterion and production.

Figure 5: Joint 95% confidence intervals for $w = [1 \ 0]^T$ (right), $w = [0.5 \ 0.5]^T$ (left) and $w = [0 \ 1]^T$ (left) for the combination ME-criterion and production.

Figure 6: Pareto front obtained by the ENNC (left) and the WS (right) for the combination ME-criterion and D-criterion.

Figure 7: Evolution substrate and biomass concentration (upper left and right), volume and controls (lower left and right) for 5 weight combinations ME-criterion and D-criterion.

Figure 8: Joint confidence intervals for $w = [1 \ 0]^T$ (right), $w = [0.5 \ 0.5]^T$ (left) and $w = [0 \ 1]^T$ (left) for the combination of the ME- and D-criterion.

Figure 9: Pareto front obtained by the NBI (left) and filtered using the Pareto filter (right) for the combination M-criterion and correlation.

Figure 10: Evolution substrate and biomass concentration (upper left and right), volume and controls (lower left and right) for 4 weight combinations of M-criterion and decorrelation.

Figure 11: Joint confidence intervals for $w = [1 \ 0]^T$ (left), $w = [0.5 \ 0.5]^T$ (left) and $w = [0 \ 1]^T$ (right) for the combination of the M-criterion and decorrelation criterion.

Figure 12: Pareto front for combining the tracking objective and the D-criterion (left) and for combining the tracking objective, correlation criterion and the D-criterion (right).

Figure 13: Evolution of the different states and control for three different weight combinations of the tracking objective and D-criterion and two additional weights combinations for the extension to incorporate the decorrelation.

Figure 14: Joint 95% confidence intervals for the three different weight combinations for the two objective case (left) and the additional two confidence intervals resulting from the three objective case (right).

Figure 15: Pareto front obtained by the NNC for the combination of A- and E-criterion.

Figure 16: Evolution of the different states and control for three different weight combinations of A- and E-criterion.

Figure 17: Joint 95% confidence intervals for the three different weight combinations for combining the A- and E-criterion.

Figure 18: Presentation of the Pareto front for 4 objectives in a bar diagram for the fed-batch bioreactor (upper) and for the Lotka Volterra fishing problem (lower).