Optimal experiment design under parametric uncertainty : a comparison of a sensitivities based approach versus a polynomial chaos based stochastic approach

Postprint version of paper published in Chemical Engineering Science 2020, 115651.
The content is identical to the published paper, but without the final typesetting by the publisher.

Optimal experiment design under parametric uncertainty: a comparison of a sensitivities based approach versus a polynomial chaos based stochastic approach

Philippe Nimmegeers[a], Satyajeet Bhonsale[a], Dries Telen[a], Jan Van Impe[a,*]

[a] *KU Leuven, Department of Chemical Engineering, BioTeC+ & OPTEC, Belgium*

Abstract

In order to estimate parameters accurately in nonlinear dynamic systems, experiments that yield a maximum of information are invaluable. Such experiments can be obtained by exploiting model-based optimal experiment design techniques, which use the current guess for the parameters. This guess can differ from the actual system. Consequently, the experiment can result in a lower information content than expected and constraints are potentially violated. In this paper an efficient approach for stochastic optimal experiment design is exploited based on polynomial chaos expansion. This stochastic approach is compared with a sensitivities based approximate robust approach which aims to exploit (higher order) derivative information. Both approaches aim at a more conservative experiment design with respect to the information content and constraint violation. Based on two simulation case studies, practical guidelines are provided on which approach is best suited for robustness with respect to information content and robustness with respect to state constraints.

Keywords: Optimal experiment design, Stochastic dynamic optimization, Fisher information matrix, Polynomial chaos expansion, Parametric uncertainty, Approximate robust optimization

1. Introduction

Performing experiments (in a (bio)chemical setting) is usually costly (Bouvin et al., 2015) as measurements have to be taken and are often analyzed manually. Furthermore, an accurate estimation of the parameters in nonlinear processes is not trivial. In order to reduce the experimental burden optimal experiment design (OED) approaches have been developed and applied in many different (bio)chemical applications (Espie and Macchietto, 1989; Asprey and Macchietto, 2002; Jauberthie et al., 2006; Cappuyns et al., 2007; Schenkendorf et al., 2009; Telen et al., 2012b, 2014). So, the main aim of optimal experiment design is to design control inputs and sampling schedules such that the experiment is as informative as possible.

*Corresponding author
Email address: jan.vanimpe@kuleuven.be (Jan Van Impe)

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

An overview of the state-of-the-art for nonlinear dynamic systems can be found in Franceschini and Macchietto (2008).

In OED an experiment is planned to estimate the model parameters, however, the model-based technique depends on these uncertain/unknown parameter values. As a result, parametric uncertainty has two consequences. First, the information obtained by performing the experiment must be ensured for all possible true system parameter values. In this work, this is called robustness with respect to information content (Asprey and Macchietto, 2000). In literature, several approaches have been explored to tackle this issue. A practical option is to iterate between the parameter estimation and subsequently compute the experiment design using the current parameter estimates, as in e.g., Walter and Pronzato (1997). Such approach is however time consuming and not necessarily robust in the sense that the experiment is ensured for all possible true system parameter values. A first approach to design robust experiments is to cast them in a max-min optimization problem (Pronzato and Walter (1988); Körkel et al. (2004); Rojas et al. (2007)). In Körkel et al. (2004) the inner optimization loop is solved explicitly with a linear approximation. Welsh and Rojas (2009) proposed a scenario-based robust experiment design approach which uses a probabilistic relaxation of the worst case robust paradigm. In this case it is considered that robustness with respect to a large majority of situations is sufficient rather than against all possible situations. The number of scenarios is set by the designer. A different approach is to compute the expected value of the scalar function of the Fisher information matrix over the parameter space if stochastic information on the parameter uncertainty is available. This idea has been introduced in Pronzato and Walter (1985) and was for the first time applied to a dynamic system for a Gaussian parameter distribution in Asprey and Macchietto (2002). In the latter work the expected value is computed by integrating numerically over the parameter space. In the frame of computing the expected value of the scalar function of the Fisher information matrix, Chu and Hahn (2008) presented an iterative approach integrating parameter set selection and optimal experiment design under uncertainty in which a genetic algorithm is used to determine the set of parameters to be estimated and a simultaneous perturbation stochastic approximation

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

computes the experimental conditions. The parameters to be estimated and the experimental conditions are the optimization variables, yielding a mixed integer nonlinear programming problem. A collection of parameter sets is returned and optimal experiment designs are computed for each of these sets. Bayesian robust experiment design is another possibility, in which preliminary data are incorporated to maximize the expected value over the prior parametric uncertainty distribution of an objective function quantifying the information content e.g., Liepe et al. (2013). Note that for the experiment design of multiple-input multiple-output systems, also a robust experiment design based on the steady state gain matrix can be used as outlined in Häggblom (2017). Although not accounting for dynamics, the reformulation of Bruwer and MacGregor (2006) made it possible to include linear input and output constraints in this approach.

A second consequence of the parametric uncertainty are the potential violations of state constraints as the model parameters differ from the *true* system parameters. So, besides robustness with respect to the information content, the optimally designed experiment has to be *robust with respect to state constraints*. These issues are related to the field of stochastic/robust optimal control. If stochastic information is available, chance constraints can be formulated (Wendt et al., 2002; Srinivasan et al., 2003; Mitra, 2009; Galvanin et al., 2010; Recker et al., 2012; Mesbah et al., 2014; Telen et al., 2015). It can be assumed that this stochastic information originates from previous parameter identifications or a literature review (Walter and Pronzato, 1997; Franceschini and Macchietto, 2008; Hjalmarsson, 2009). In a different set-up the parameters can be considered to lie within a given compact set. In this case, it is desirable to guarantee that all constraints are satisfied in all possible worst case situations and/or to know what is the possible performance loss. The work of Houska et al. (2012) presents an approach for nonlinear optimal control which guarantees to be robust if the uncertainties are bounded. In Telen et al. (2013a), this approach is extended for optimal experiment design.

The definition of the expected value entails the computation of a multidimensional integral over the parameter space of the scalar function of the Fisher information

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

matrix (Pronzato and Walter, 1985). In real (bio)chemical applications these multidimensional integrals can often not be evaluated analytically as a function of the decision variables and therefore they need to be approximated with a sampling-based method, e.g. Gauss quadrature or Monte Carlo sampling (Asprey and Macchietto, 2002),(Debusschere et al., 2004). Therefore, instead of computing this integral, the expected value is approximated in this article using polynomial chaos expansion (PCE) (Wiener, 1938). The main advantage of polynomial chaos expansion over similar techniques as the unscented transformation or sigma point approach (Julier and Uhlmann, 1996; Kawohl et al., 2007; Telen et al., 2014) is its applicability to non-symmetric parametric uncertainty distributions (Wiener, 1938), (Xiu and Karniadakis, 2002), while the unscented transformation is restricted to symmetric, unimodal distributions. The basic idea of PCE is to approximate a function by a polynomial depending on the uncertain parameters. The coefficients of this polynomial can subsequently be used to compute the statistical moments as the expected value and variance (Nagy and Braatz, 2007; Mesbah et al., 2014). These statistical moments can be used for the objective or constraint functions and hence allow for a more robust probabilistic problem formulation (Galvanin et al., 2010). Recently, a novel arbitrary polynomial chaos expansion algorithm has been presented in Paulson et al. (2017), which does not require prior knowledge on the parametric uncertainty distribution but computes the orthogonal polynomial basis functions based on data (i.e., raw moments of the random variables).

In summary, model-based optimal experiment techniques can be used to design experiments that yield a maximum of information to estimate parameters accurately in nonlinear dynamic systems. These techniques, however, use a current guess of the parameters which can be different from the actual system. Consequently, the experiment can result in a lower information content than expected and constraints are potentially violated. As mentioned above, different optimal experiment design techniques exist to design experiments which are robust with respect to information content (optimality) and robust with respect to constraint violations (feasibility). The overall goal of this paper is to study and compare two optimal experiment design approaches which can be used to design robust experiments: a sensitvities-

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

based approximate robust approach (originating from a robust min-max optimal experiment design formulation) and a polynomial chaos expansion based stochastic approach. Based on two simulation case studies, practical guidelines are provided on which approach is best suited for (i) robustness with respect to information content and (ii) robustness with respect to state constraints. The assessment of the different OED approaches is based on the information content (i.e., the OED objective function value), the number of constraint violations and the computational (CPU) time.

This paper is structured as follows. In Section 2 the mathematical formulation of OED, robust OED and expected value OED are introduced. In Section 3 the actual OED optimization problems are presented, i.e., the approximate robust approach of Körkel et al. (2004) and the PCE-based stochastic approach. Section 4 introduces the case studies and describes the obtained numerical results. Section 5 summarizes the main conclusions of this paper.

2. Mathematical formulations

This section is structured as follows. First, OED is presented as an optimization problem for nonlinear dynamic systems. The adaptations to the standard OED formulation in order to obtain a robust or a stochastic approach are presented in subsections 2.2 and 2.3.

2.1. Optimal experiment design for dynamic systems

Optimal experiment design for parameter estimation (OED-PE) is used to design experiments that reduce the variance on the parameter estimates. The objective function used in OED is a scalar function of the parameter estimation variance-covariance matrix. Different techniques exist to compute the parameter estimation variance-covariance matrix and a brief overview is presented below.

A first technique is based on the Fisher information matrix (FIM). The inverse of the Fisher information matrix approximates the Cramér-Rao bound, a measure for the lower bound on the variance of estimators, assuming unbiased estimators (Ljung, 1999), (Walter and Pronzato, 1997). This is the most common technique

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

and the technique used throughout this article.

Other methods exist to approximate the parameter estimation variance-covariance matrix: Telen et al. (2013b) proposed a technique based on the solution of a Riccati differential equation that allows to directly account for process noise and requires a lower number of differential states than the Fisher information matrix approach.

The techniques of Heine et al. (2008) and Schenkendorf et al. (2009) both rely on the sigma point method/unscented transformation which approximates a distribution with a fixed number of parameters, the sigma points. The method presented by Heine et al. (2008) uses a derivative free filter based on a polynomial interpolation with a maximum a posteriori update by a Bayesian formulation to compute the parameter estimation variance-covariance matrix. The method presented by Schenkendorf et al. (2009) uses the sigma points to sample from the measurement error distribution and add these errors to the output profiles for the current best guess of the parameter values. This results in $2n_y+1$ measurement profiles on which subsequently a separate parameter estimation procedure has to be performed. These $2n_y + 1$ parameter sets are then used to compute the expected value of the parameters and parameter estimation variance-covariance matrix.

Monte Carlo simulations can also be used to obtain an empirical estimate of the parameter distribution by simulating N realizations from the noise distribution, and performing parameter estimation for each of the obtained datasets. This is computationally inefficient as many realizations have to be taken to obtain sufficiently accurate parameter estimation variance-covariance matrix computations, e.g.: 500 realizations in Balsa-Canto et al. (2008) and 10000 realizations in Schenkendorf et al. (2009).

As in this article the Fisher information matrix method for computing the parameter estimation variance-covariance matrix is used, the mathematical problem formulation with this method is introduced.

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

The complete *classic, dynamic OED problem formulation* incorporating the required sensitivities and the Fisher information matrix is in this paper considered as follows (Telen et al., 2014):

$$\min_{u(\cdot),x(\cdot),F(\cdot)} \Phi(F(t_\mathrm{f})) \tag{1}$$

subject to:

$$\frac{dx}{dt}(t) = f(x(t),u(t),p,t) \quad \text{with} \quad x(0)=x_0, \tag{2}$$

$$\frac{d}{dt}\frac{\partial x}{\partial p}(t) = \frac{\partial f}{\partial x}\frac{\partial x}{\partial p}(t)+\frac{\partial f}{\partial p} \quad \text{with} \quad \frac{\partial x}{\partial p}(0)=\frac{\partial x_0}{\partial p}, \tag{3}$$

$$\frac{d}{dt}F(t) = w(t)\frac{\partial x}{\partial p}(t)^\top \frac{dh(x(t))}{dx}^\top Q(t)^{-1}\frac{dh(x(t))}{dx}\frac{\partial x}{\partial p}(t) \quad \text{with} \quad F(0)=0, \tag{4}$$

$$0 \geq c_\mathrm{i}(x(t),u(t),t), \tag{5}$$

The first equation denotes the objective function, which is in this article a scalar function $\Phi(\cdot)$ of the Fisher information matrix. Typically, this is one of the alphabetic criteria, i.e., A- (minimize trace of the inverse of the Fisher information matrix), D- (maximize determinant of the Fisher information matrix) or E-criterion (maximizing the smallest eigenvalue of the Fisher information matrix) (Walter and Pronzato, 1997). Equation (2) describes the actual system dynamics with the states $x(t) \in \mathbb{R}^{n_\mathrm{x}}$, the controls $u(t) \in \mathbb{R}^{n_\mathrm{u}}$ and the parameters $p \in \mathbb{R}^{n_\mathrm{p}}$. These parameters are time-invariant but an experiment to determine their exact values based on measurements is required. Equations (3) and (4) are the required sensitivity equations and the continuous formulation of the Fisher information matrix. Equation (3) requires the solution of $n_\mathrm{p}n_\mathrm{x}$ additional ordinary differential equations. Computing Equation (4) yields the Fisher information matrix. Therefore, the objective function which represents the total information content is evaluated at t_f, the final time. Here the function $h(x(t))$ denotes the measurement function which can depend non-linearly on the states $x(t)$, $w(t) \in [0,1]$ is a function indicating whether a sample is taken (it is a relaxed function, avoiding that a mixed-integer optimization problem needs to be solved) and $Q(t)$ denotes the measurement variance-covariance matrix. Without loss of generality, these can also be computed based on a summation depending whether a discrete or a continuous measurement frame is employed. The symmetry in $F(t)$ can be exploited to reduce the number of ordinary differential

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

equations (and hence the number of states), i.e., $\frac{n_\mathrm{p}}{2}(n_\mathrm{p}+1)$ instead of n_p^2. Equation (5) denotes the present constraints $c_\mathrm{i} \in \mathbb{R}^{n_\mathrm{c}}$. Consequently, the total number of states involved in OED (n_OED) equals:

$$n_\mathrm{OED} = n_\mathrm{x} + n_\mathrm{p} \cdot n_\mathrm{x} + \frac{n_\mathrm{p}}{2} \cdot (n_\mathrm{p}+1)\ . \tag{6}$$

2.2. Robust optimal experiment design

Assume that the parameters p are normally distributed, with nominal parameter value (mean value) p_nom and variance Σ . With a confidence quantile γ, the following ellipsoidal joint confidence region for the model parameters can be considered:

$$\|p - p_\mathrm{nom}\|^2_{\Sigma^{-1}} \leq \gamma, \tag{7}$$

with the norm $\|p\|_{\Sigma^{-1}} = \left(p^\top \Sigma^{-1} p\right)^{(1/2)}$.

Assuming that the parametric uncertainty is characterized by a normal distribution, the sum of squared parameter estimation errors,

$$\|p - p_\mathrm{nom}\|^2_{\Sigma^{-1}} = (p - p_\mathrm{nom})^\top \Sigma^{-1} (p - p_\mathrm{nom})\,, \tag{8}$$

is $\chi^2(n_\mathrm{p})$ distributed, the objective and constraint functions in Equations (1) and (5) can be replaced by the following equations in a *robust, dynamic OED problem formulation* (Körkel et al., 2004):

$$\min_{u(\cdot),x(\cdot),F(\cdot)} \max_{\|p-p_\mathrm{nom}\|^2_{\Sigma^{-1}} \leq \gamma} \Phi(F(t_\mathrm{f})) \tag{9}$$

$$0 \geq \max_{\|p-p_\mathrm{nom}\|^2_{\Sigma^{-1}} \leq \gamma} c_i(x(t),u(t),t) \quad i = 1,\ldots,n_\mathrm{c} \tag{10}$$

Note that the problem formulation in Equations (9)-(10) is a conventional worst-case approach as in e.g., Pronzato and Walter (1988). Contrary to standard robust approaches, it is assumed in this article that the parameters can be described by a known uncertainty distribution. To guarantee a solution to the inner maximization problem for a closed set of model parameters, the sum of squared parameter

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

estimation errors is limited to a certain preset quantile γ.

2.3. Stochastic optimal experiment design

Another approach to account for parametric uncertainty in optimal experiment design is stochastic optimal experiment design. Stochastic optimization approaches exploit knowledge on a known probability distribution of the uncertainty to formulate expected values of the model responses, as e.g., the objective function, and to formulate chance constraints (Nagy and Braatz, 2004). In this article, single chance constraints are considered. The parametric uncertainty distribution (or at least information on the moments) is propagated through the (nonlinear) dynamic system to approximate the statistical moments (e.g., expected value and variance) of the model's states or response functions (e.g., objective function, outputs, constraint functions). Furthermore, chance constraints express that the probability of a constraint to be violated is smaller than or equal to a preset probability ϵ_i (Wendt et al., 2002), (Mesbah and Streif, 2015):

$$\epsilon_\mathrm{i} \geq \mathbf{Pr}\left[0 < c_\mathrm{i}(x(t), u(t), t)\right] \tag{11}$$

The preset probability ϵ_i is set based on how much constraint violations are acceptable, the more critical the constraint, the lower the probability ϵ_i is set. In this article ϵ_i is set equal to 5%.

Stochastic optimization approaches exploit knowledge on a known probability distribution of the uncertainty to formulate expected values of the model responses, as e.g., the objective function, and to formulate chance constraints (Nagy and Braatz, 2004). In this dissertation, single chance constraints are considered. Single chance constraints express that the probability of a constraint to be violated is smaller than or equal to a preset probability ϵ_i (Wendt et al., 2002).

In a *stochastic, expected value dynamic OED problem formulation with chance constraints* the objective function in Equation (1) and constraint functions in Equation (5), are replaced by Equations (12) and (13), respectively.

$$\min_{u(\cdot), x(\cdot), F(\cdot)} \mathbb{E}\left[\Phi(F(t_\mathrm{f}))\right] \tag{12}$$

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

subject to:

$$\epsilon_{\mathrm{i}} \geq \mathbf{Pr}\left[0 \leq c_{\mathrm{i}}(x(t), u(t), t)\right] \quad i = 1, \ldots, n_{\mathrm{c}} \tag{13}$$

Note that similarly to Asprey and Macchietto (2002) an expected value is used in the objective function and this formulation ensures that the system is kept within a feasible region with specified probability as in e.g., Galvanin et al. (2010).

3. Reformulation to the actual OED problems

In this section, the approximate robust OED formulation is presented first in which the inner maximization problem is linearized. Subsequently, polynomial chaos expansion is applied to stochastic OED as in Mesbah and Streif (2015), Nimmegeers et al. (2017). Finally, the approximate robust and PCE based stochastic OED formulations are compared.

3.1. Sensitivities based approximate robust OED reformulation

The approach of Körkel et al. (2004) consists of calculating a first order Taylor series approximation of the objective function, which transforms the inner non-convex maximization problem to a convex maximization of a linear function (i.e., $\Phi(F(t_{\mathrm{f}})) + \frac{d}{dp}\Phi(F(t_{\mathrm{f}}))(p - p_{\mathrm{nom}})$) subject to a convex quadratic constraint (i.e., $\|p - p_{\mathrm{nom}}\|^2_{\Sigma^{-1}} \leq \gamma$). By taking these assumptions, the inner maximization problem has the following solution (as derived in Appendix A) in contrast with what has been derived in (Körkel et al., 2004):

$$\Phi(F(t_{\mathrm{f}})) + \sqrt{\gamma}\left\| \frac{d}{dp}\Phi(F(t_{\mathrm{f}}))\right\|_{\Sigma} \tag{14}$$

Although the evaluation of this solution to the inner maximization problem seems straightforward, the implementation of the derivative of the Fisher information matrix with respect to the parameters is needed to compute $\frac{d}{dp}\Phi(F(t_{\mathrm{f}}))$ in the objective function of the approximate robust OED problem formulation. Different mathematical approaches exist to implement the computation of the derivative of the Fisher information matrix elements with respect to the parameters as for instance, finite differences or calculating second order sensitivities through tensor

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

variational equations (Vassiliadis et al., 1999), (Balsa-Canto et al., 2001), (Telen et al., 2012a). Therefore this method is referred to in this article as a sensitivities based approximate robust approach. Moreover, advanced automatic differentiation tools as e.g., `casADi` (Andersson et al., 2012) can be exploited to retrieve the Jacobian of the Fisher information matrix efficiently without the need for additional states. This last approach is followed in this paper. Note that $\sqrt{\gamma}\left\|\frac{d}{dp}\Phi(F(t_{\mathrm{f}}))\right\|_{\Sigma}$ can be seen as an approximation of the standard deviation on the OED objective function $\Phi(F(t_{\mathrm{f}}))$.

The same approach can be followed for the constraint function, i.e., the constraint should be satisfied in the worst case as shown in Equation (15).

$$0 \quad \geq \max_{\|p-p_{\mathrm{nom}}\|_{\Sigma^{-1}}} c_i(x_{\mathrm{i}}(t), u(t), t) \quad i = 1, \ldots, n_{\mathrm{c}}. \tag{15}$$

Similarly as for the objective function, a first order Taylor series approximation of the constraint function can be made, resulting in a convex maximization of a linear function (in this case $c_i(x_{\mathrm{i}}(t), u(t), t)$), subject to a convex quadratic constraint (i.e., $\|p - p_{\mathrm{nom}}\|_{\Sigma^{-1}}$. This results in the following constraint:

$$0 \quad \geq \quad c_{\mathrm{i}}(x(t), u(t), t) + \sqrt{\gamma}\left\|\frac{d}{dp}c_{\mathrm{i}}(x(t), u(t), t)\right\|_{\Sigma} \tag{16}$$

The norm $\left\|\frac{d}{dp}c(x(t), u(t), t)\right\|_{\Sigma}$ can be seen as an approximation of the standard deviation on the constraint function $c_{\mathrm{i}}(x(t), u(t), t)$. Note that the required derivative of the constraint function with respect to the parameters $\frac{d}{dp}c_{\mathrm{i}}(x(t), u(t), t)$ can be easily computed from the sensitivity states.

$$\left\|\frac{d}{dp}c_{\mathrm{i}}(x(t), u(t), t)\right\|_{2,\Sigma} = \sqrt{\left(\frac{dx}{dp}\right)^{\top}\left(\frac{dc_{\mathrm{i}}}{dx}\right)^{\top}\Sigma\frac{dc_{\mathrm{i}}}{dx}\frac{dx}{dp}} \tag{17}$$

Equation (17) equals the first order approximation of the constraint function's variance-covariance matrix (Nagy and Braatz, 2004), (Telen et al., 2015).

Hence, for the sensitivities based approximate robust OED formulation the following objective function and constraint function can be used to replace Equations

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

(1) and (5) in the formulation of the general OED problem (Equations (1)-(2)):

$$\Phi_{\mathrm{rob}}(F(t_{\mathrm{f}})) = \Phi(F(t_{\mathrm{f}})) + \sqrt{\gamma}\left\| \frac{d}{dp}\Phi(F(t_{\mathrm{f}}))\right\|_{\Sigma} \tag{18}$$

subject to:

$$0 \;\geq\; c_{\mathrm{i}}(x(t), u(t), t) + \sqrt{\gamma}\sqrt{\left(\frac{dx}{dp}\right)^{\top}\left(\frac{dc_{\mathrm{i}}}{dx}\right)^{\top}\Sigma\frac{dc_i}{dx}\frac{dx}{dp}}, \quad i = 1, \ldots, n_{\mathrm{c}} \tag{19}$$

3.2. Polynomial chaos based stochastic OED formulation

In stochastic optimal experiment design, the constraints can be formulated as chance constraints. However, addressing these chance constraints in dynamic optimization is computationally challenging as pointed out in e.g., Mesbah et al. (2014). Cantelli-Chebyshev's inequality can be used to reformulate these chance constraints as the following equivalent deterministic constraints (Mesbah and Streif, 2015):

$$0 \geq \mathbb{E}\left[c_{\mathrm{i}}\right] + \alpha_{c_{\mathrm{i}}}\sqrt{\mathbb{V}\mathrm{ar}\left[c_{\mathrm{i}}\right]} \tag{20}$$

In Equation (20), $\mathbb{E}\left[c_{\mathrm{i}}\right]$ and $\mathbb{V}\mathrm{ar}\left[c_{\mathrm{i}}\right]$ express the expected value and variance of the constraint function c_{i}, respectively. The coefficient $\alpha_{c_{\mathrm{i}}}$ is introduced as a *backoff parameter* (e.g., (Galvanin et al., 2010)) and can be seen as an uncertainty quantile (Telen et al., 2015). Note that the objective function can also include a penalization term for large variations by adding a term accounting for the variance weighted with a backoff parameter:

$$\mathbb{E}\left[J\right] + \alpha_J\sqrt{\mathbb{V}\mathrm{ar}\left[J\right]} \tag{21}$$

Polynomial chaos expansion (PCE) can be used for the computation of the variance and expected value of model responses (e.g., objective function, constraint function, etc.). Contary to other similar uncertainty propagation techniques as the unscented transformation or sigma point approach (Julier and Uhlmann, 1996; Kawohl et al., 2007; Telen et al., 2014), PCE is not limited to symmetric, unimodal distributions but can also be applied to non-symmetric parametric uncertainty distributions (Wiener, 1938), (Xiu and Karniadakis, 2002). The rationale of polynomial chaos

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

expansion is to approximate the model response (e.g., objective function, constraint function, etc.) as a sum of orthogonal polynomials (i.e., polynomials of which the inner product equals zero) through *PCE collocation points.* These polynomials are a function of the uncertain variable for which a probability distribution is assumed to be given (Mesbah and Streif, 2015),(Nimmegeers et al., 2016).

Consider the d-th order polynomial chaos expansion of the OED objective function $\Phi(F(t_{\mathrm{f}}))$, with a given distribution for the parameters p (with a given expectation value $\bar{p}$ and variance-covariance matrix P_{pp}) is defined in Equation (22):

$$\Phi(F(t_{\mathrm{f}})) \approx \sum_{\mathrm{j}=0}^{L-1} a^{d}_{\Phi,\mathrm{j}} \Psi_{\mathrm{j}}(p). \tag{22}$$

Here PCE is formulated using a term based index j ($\mathrm{j} = 0, \ldots, L-1$). The symbol $a^{d}_{\Phi,\mathrm{j}}$ denotes the unknown PCE coefficients and $\Psi_{\mathrm{j}}(y)$ the multivariate orthogonal polynomials. The total number of terms L in the polynomial chaos expansion of order d depends on the number of uncertain variables n and the order of the expansion d:

$$L = \frac{(n+d)!}{n!d!}. \tag{23}$$

Intrusive and *non-intrusive* methods exist to estimate the unknown coefficients $a^{d}_{\Phi,\mathrm{j}}$. This distinction is based on the extent to which the problem needs to be reformulated. More specifically, *intrusive methods* develop a deterministic set of equations for the coefficients $a^{d}_{\Phi,\mathrm{j}}$ based on a Galerkin projection of the approximation error between the model response function (for instance the OED objective function $\Phi(F(t_{\mathrm{f}}))$) and its polynomial chaos expansion. Note that for *intrusive* methods the model response needs to be explicitly known and preferably the explicit model response function is a polynomial function. In *non-intrusive methods* the model is considered as a black box and exact expressions for the model response are not required. All *non-intrusive* methods can be considered as a weighted sum of model response evaluations in n_{s} sampling points.

In this work a non-intrusive PCE method based on least squares regression is followed in order to determine the unknown coefficients $a^{d}_{\Phi,\mathrm{j}}$. The model is evaluated

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

in *sampling points*, which are selected from the roots of the higher order (i.e., $d+1$) orthogonal polynomial for each uncertain parameter. For more details on the computation of the PCE coefficients with this least squares regression approach, the reader is referred to Nimmegeers et al. (2016).

In summary, the PCE coefficients are computed as a weighting of the function $\Phi(F(t_\mathrm{f}))$ evaluated at the different *sampling points* π_i. The objective function and constraints for the PCE based stochastic OED formulation are defined as:

$$\Phi_\mathrm{PCE} = a^d_{\Phi,0} + \alpha_\Phi \sqrt{\sum_{j=1}^{L-1} \left(a^d_{\Phi,\mathrm{j}}\right)^2 \mathbb{E}\left[\Psi^2_\mathrm{j}(p)\right]} \tag{24}$$

subject to:

$$0 \quad \geq \quad a^d_{\mathrm{c_i},0} + \alpha_{c_i} \sqrt{\sum_{\mathrm{j}=1}^{L-1} \left(a^d_{c_i,\mathrm{j}}\right)^2 \mathbb{E}\left[\Psi^2_\mathrm{j}(p)\right]} \tag{25}$$

where $\mathbb{E}\left[\Psi^2_\mathrm{j}(p)\right]$ is computed offline.

3.3. Comparison of the OED formulations

In Table 1 the objective function and constraint formulations are shown for the nominal (not accounting for uncertainty) optimal experiment design, sensitivities based approximate robust experiment design and the PCE based stochastic experiment design approaches. From Table 1 it can be seen that the approximate robust and PCE based stochastic OED approaches formulate the objective (or constraint) function as a sum of two terms in which the second term is an approximation of the variance on the objective (or constraint) function, weighted with a backoff parameter.

The major difference between the approximate robust OED formulation and the PCE based stochastic OED formulation is the number of required states. In the approximate robust OED formulation, the model is only evaluated in the nominal parameter values. However, depending on the approach used for the evaluation of the derivative of the Fisher information matrix with respect to the parameters, the number of states differs. If tensor variational equations (Vassiliadis et al., 1999), (Balsa-Canto et al., 2001), (Telen et al., 2012a) are used, the number of states corresponds to $n_\mathrm{rob,tensor-approx} = n_\mathrm{x} + (n_\mathrm{p}+1)n_\mathrm{x}n_\mathrm{p} + (n_\mathrm{x}n_\mathrm{p}+1)n_\mathrm{p}(n_\mathrm{p}+1)/2$. In

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

case that an automatic differentiation tool as `casADi` is used, the number of states corresponds with $n_{\text{rob,approx}} = n_{\text{OED}} = n_{\text{x}} + n_{\text{x}} n_{\text{p}} + n_{\text{p}}(n_{\text{p}}+1)/2$. However, in the PCE based stochastic OED formulation the model is evaluated in the n_{s} sampling points, leading to a system of $n_{\text{PCE}} = n_{\text{s}}(n_{\text{x}}+n_{\text{x}}n_{\text{p}}+n_{\text{p}}(n_{\text{p}}+1))/2$ states, which are much easier parallellized as they consist of copies of the same system only differing in the model parameters.

In the approximate robust OED formulation the worst-case objective function is computed by a linearization of the inner maximization problem. The computation of this worst-case objective function is related to the assumption of a normal distribution of the parametric uncertainty (hence a chi-square distribution of $\|p - p_{\text{nom}}\|^2_{\Sigma^{-1}}$). This leads to two terms in which one term is the objective function evaluated in the nominal parameter values and the second term contains the first order approximation of the variance on the objective function, weighted with the square root of a chi-square confidence quantile.

In the PCE based stochastic OED formulation, an expected value objective function is formulated based on the parametric uncertainty distribution. To penalize for the variance on the objective function, a variance-related term can be added to the objective function, weighted with a backoff parameter. These terms are both based on the computation of a weighted sum of the objective function evaluated in the different sampling points.

Besides the difference in practical computation of these terms, the underlying reasoning is different for both methods. Similarly to the objective function, worst case constraint functions are computed in the approximate robust OED formulation. In the PCE based stochastic OED formulation, chance constraints are considered expressing that the probability of a constraint to be violated is smaller than or equal to a certain value.

A final difference between the two formulations is the choice of the backoff parameters. In the approximate robust OED formulation these backoff parameters are based on the assumption that the sum of squared parameter estimation errors is chi-square distributed and γ corresponds to a chi-square quantile. For the PCE based stochastic approaches the choice of this parameter can be related to a quantile (if the distribution of the considered response (i.e., objective function or constraint

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

function) is known) or based on Cantelli-Chebyshev's inequality (Mesbah and Streif, 2015). Telen et al. (2015) presents an iterative strategy for selecting this backoff parameter.

4. Results

Two case studies are investigated in this work. The first case study is a Lotka Volterra predator prey model augmented with a fishing term. In the second case study the jacketed tubular reactor is considered. In both case studies information optimality of the experiment design is studied. As a reactor temperature state constraint is present in the second case study, the feasibility of the experiment design (in terms of constraint violations) is also studied more in depth in the second case study.

From the formulation in (1)-(5), it is clear that OED is a type of dynamic optimization problems. In dynamic optimization an optimal value for the control inputs has to be found for every $t \in [0, t_\mathrm{f}]$. OED for nonlinear dynamic models is a subclass of dynamic optimization which quickly leads to a high number of states. These problems are solved in this work by discretizing the controls via single shooting using `casADi` (Andersson et al., 2012). The resulting NLP is solved with `IPOPT` (Wächter and Biegler, 2006).

Before starting with the case studies, firstly the indicators that are used for the assessment of the different OED approaches are introduced.

4.1. Assessment of the different OED approaches

The performance of the different OED approaches is assessed in terms of optimality (information content), feasibility (constraint violations) and computational time. In this article two metrics are used for the information content: the E-criterion and the D-criterion.

The E-criterion aims at minimizing the largest eigenvalue of the variance-covariance matrix. Using the Fisher information matrix approach for OED, this corresponds to maximizing the smallest eigenvalue of the Fisher information matrix. Geometrically, an E-optimal design minimizes the length of the largest axis of the joint

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

confidence region (Kiefer and Wolfowitz, 1959). Hence, the greater the smallest eigenvalue of the Fisher information matrix, the higher the information content. This criterion is used in the first case study, the Lotka Volterra fishing problem.

The D-criterion minimizes the determinant of the variance-covariance matrix and is implemented in this article as the maximization of the determinant of the Fisher information matrix. A D-optimal design minimizes the volume of the confidence region (Kiefer and Wolfowitz, 1959). Hence a high determinant of the Fisher information matrix corresponds with a high information content. This criterion is used in the second case study, the jacketed tubular reactor.
In order to assess the performance of the OED approaches Monte Carlo simulations have been executed in which parameter values are randomly taken from the parametric uncertainty distribution to simulate the system with the computed optimal experimental inputs. The E-criterion values (for the first case study) abd D-criterion values (for the second case study) are evaluated and compared for the different OED approaches.

In the second case study, a reactor temperature state constraint is present and the feasibility of the experiment design (in terms of constraint violations) is also studied by means of Monte Carlo simulations. The lower the number of constraint violations the more robust it is with respect to constraint violations.

Note that two parameters are typically set by the user; α for the PCE-based stochastic approach and γ for the approximate robust approach. In the first case study, emphasis is on robustness with respect to information content and as no state constraints are present, robustness with respect to constraint violations is not studied. In the first case study α and γ are selected based on quantiles as mentioned in subsection 4.2. In the second case study, emphasis is on robustness with respect to constraint violations due to the reactor temperature state constraint. In this case study α and γ are seen as backoff parameters and as outlined by Telen et al. (2015) to reduce the number of constraint violations.

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

4.2. A Lotka Volterra fishing problem - robustness in information content

In this first case study a Lotka Volterra fishing problem (Sager, 2013; Telen et al., 2012b) is considered. The goal of this model is to track a predetermined steady state value for both the predator and prey states where typically the decision to fish is considered to be binary. In the implementation of this case study, the problem is solved in a relaxed version, i.e., $u \in [0, 1]$ and the strategy for connecting the optimal control values to binary values from Sager et al. (2009) is applied. Two fish populations live in a pond: a prey and a predator population. In this case study the aim is to develop an optimal fishing strategy $u(t)$ and sampling strategy $w(t)$ (i.e., the population measurement by the diver) to estimate the parameters in the prey and predator mass balances related to the interaction between predator and prey.

The model equations are:

$$\frac{dx_1}{dt} = x_1 - p_1 x_1 x_2 - 0.4 x_1 u, \tag{26}$$

$$\frac{dx_2}{dt} = -x_2 + p_2 x_1 x_2 - 0.2 x_2 u, \tag{27}$$

where x_1 is the biomass of the prey and x_2 the biomass of the predator. The symbol u is the fishing control. The initial conditions are set to: $x_1(0) = 0.5$ and $x_2(0) = 0.7$, furthermore the final time is fixed at $t_f = 12$. The assumed mean parameter values are $p_1 = 1$ and $p_2 = 1$.

Both states are considered to be measurable. The parameter variance-covariance matrix is assumed to be

$$V = \begin{pmatrix} 0.01 & 0 \\ 0 & 0.01 \end{pmatrix}. \tag{28}$$

Remark: If the parameter distribution is not known (as is often the case), an assumption can be made regarding the parameter distribution, potentially based on available experimental data and from a parameter estimation procedure and distribution fitting (often a normal distribution) or a conservative distribution as e.g., uniform distribution can be taken. If V contains correlation between the

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

parameters, then this can be accounted for in defining the sampling points by using arbitrary polynomial chaos (Paulson et al., 2017). Larger values in V elements, result in greater uncertainty and hence more conservative experiment designs.

For this first case study, only the robustness with respect to the information content is investigated as there are no critical state constraints which could lead to an infeasible situation of the system. Similarly, as for constraints, the variance with respect to the information content can also be taken into account in the stochastic optimal experiment design approaches by considering a backoff parameter α as in Equation (21).

Furthermore, the number of measurements which is allowed to be taken is constrained to 6 time units. This is motivated by experimental practice where the decision when to sample is usually one of the degrees of freedom in the experiment. The goal in this case study is to maximize the information content as expressed by the minimum eigenvalue of the Fisher information matrix. This sampling strategy $w(t) \in [0, 1]$ is implemented in a relaxed form instead of considering it as a binary decision variable and enters the OED system in the ODE for the Fisher information matrix:

$$\frac{dF(t)}{dt} = w(t) \left(\frac{\partial x}{\partial p}(t)\right)^{\intercal} \left(\frac{dh(x(t))}{dx}\right)^{\intercal} \mathbf{Q}^{-1} \frac{dh(x(t))}{dx} \frac{\partial x}{\partial p}(t) \tag{29}$$

Three scenarios have been studied in this case study to investigate the influence of accounting for the variance on the information content during the experiment design: nominal OED, PCE based stochastic OED (with expected value ED, i.e., $\alpha = 0$ and $\alpha = 1.65$) and an approximate robust design in which a 95% confidence region is considered (i.e., $\gamma = 6$). In summary, the values for α and γ have been selected in this case study as follows: $\alpha = 0$ corresponds with an expected value approach, not accounting for the variance on the OED objective function, $\alpha = 1.65$ corresponds with a 95% normal quantile taken from the OED objective function for the stochastic approach while $\gamma = 6$ corresponds with 95% chi-square quantile in the approximate robust OED approach.

Note that for this case study as well a normal as a uniform parametric uncer-

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

tainty distribution are considered for the parameters p_1 and p_2. Therefore, next to a first order (PCE1) and second order (PCE2) polynomial chaos expansion based on a normal parametric uncertainty distribution of p_1 and p_2, a second order polynomial chaos expansion has been derived based on a uniform parametric uncertainty distribution (PCE2 Uniform). To illustrate the difference between the implemented strategies, the control profiles for $\alpha_\Phi = 0$ and $\alpha_\Phi = 1.65$, i.e., the fishing control $u(t)$ and the sampling action $w(t)$, are depicted in Figure 1 (a,b,c,d) and Figure 1 (e,f,g,h), respectively. The profiles for a second order polynomial chaos expansion derived from a normal parametric uncertainty distribution (PCE2) and a second order polynomial chaos expansion derived from a uniform parametric uncertainty distribution (PCE2 Uniform) are shown in Figure 1(c,d) and Figure 1(g,h) for $\alpha_\Phi = 0$ and $\alpha_\Phi = 1.65$, respectively. For $\alpha_\Phi = 1.65$, both $u(t)$ and $w(t)$ profiles differ substantially.

4.2.1. Information content

The information content as measured by the smallest eigenvalue (i.e., E-criterion value) using the current best estimate for the parameters for $\alpha = 0$ and $\alpha = 1.65$ are presented in Table 2. Thus when the parameters of the system would be exact, there is a slight loss in information content (i.e., decrease in E-criterion value as indicated in Table 2) when using the stochastic approach compared with the nominal case of approximately 5% (PCE2 approaches) and 10% (PCE1) for $\alpha = 0$ and a loss in information content of approximately 32% (PCE1), 3% (PCE2), 13% (PCE2 uniform) for $\alpha = 1.65$. The loss in information content when comparing the approximate robust approach with the nominal case is dramatic (approximately 80%). Evaluation of the norm $\left\| \frac{d}{dp}\Phi(F(t_\mathrm{f})) \right\|_\Sigma$ for the different approaches revealed that the approximate robust approach results in the smallest norm (i.e., 4.41 for the approximate robust approach versus 16.09 in the nominal case). Since the approximate robust approach only considers the norm $\left\| \frac{d}{dp}\Phi(F(t_\mathrm{f})) \right\|_\Sigma$ evaluated at the nominal parameter values this approach results in a large backoff and dramatically low information content when compared to the other approaches.

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

4.2.2. Robustness in information content with respect to parametric uncertainty

In order to investigate the robustness of the designed experiments with respect to the parameter influence 1000 parameter realizations are drawn from the assumed normal/uniform distribution with the aforementioned mean and variance values. Subsequently the mean smallest eigenvalue and quartiles are reported for $\alpha = 1.65$ a trade-off between information content and spread of the information content (i.e., how close the values of the smallest eigenvalue are for the different parameter realizations) is made and the results are different: only the stochastic PCE2 approaches yield a higher information content as can be observed in the mean values and quartiles in Table 3, respectively. The spread is generally lower for the stochastic approach than for the nominal approaches. The approximate robust approaches result in a very low information content, but also a very low spread on the information content. A possible explanation for this very low information content, but very low spread on the information content for the approximate robust approach lies in the linearization which holds when the uncertainty is small compared to the model curvature such that higher order terms can be neglected. Depending on the case study, it can be different. Therefore this result cannot be generalized. Comparing this with the nominal and stochastic approaches, it is concluded that the approximate robust designs are too conservative (approximately 4 times lower than the nominal approaches).

The effect of the stochastic approach on the cost surface (i.e., the surface constructed by plotting the E-criterion value versus the parameter values) is visualized in Figure 2. In the neighborhood of the nominal parameter values, the nominal design outperforms the stochastic approach, however, there is a distinct region where the information content drops sharply for the nominal design while this totally absent in the stochastic approach. This exemplifies the goal of the stochastic optimal experiment design approach, i.e., the attempt to remain informative for a wide range of actual parameter realizations.

The surfaces obtained in Figure 2(a) and 2(b) are also projected in the 2D figures in Figure 3. Here the dependency in each of the different parameters is depicted. For

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

parameter p_1, it is evident in Figure 3(a) and 3(c) that on average the stochastic approach performs better than the nominal design. In Figure 3(b) and 3(d), the difference is less pronounced, however, there is a distinct area where the stochastic approach outperforms the nominal design. Note also the strong dependency of the information content on parameter p_1 in Figure 3(a) and 3(c). To conclude, the variance on the information content is lower in case $\alpha = 1.65$ (as can be observed in Figure 3) and that this comes at the cost of a reduction in overall information content when compared to $\alpha = 0$ (as can be observed in Figure 2).

4.2.3. Computation times

A final aspect in which the nominal, approximate robust and stochastic approaches are evaluated is computation time (see Table 4). This computation time is closely related to the number of states in the considered OED approach. For instance, it can be expected that the PCE2 approaches require a higher computation time (3869.12 s) than the other approaches due to the higher number of states involved in the system, i.e., six times the number of states in the nominal case. For the PCE1 approaches the computation time is higher than for the nominal approach (799.93 s), since three times the number of nominal states are evaluated. The approximate robust approach on the other hand will need a higher computation time than the nominal approach due to the additional effort in automatic differentiation that is required for the computation (553.97 s) of $\frac{d}{dp}\Phi(F(t_{\mathrm{f}}))$. The nominal OED approach only requires 91.06 s.

4.3. A jacketed tubular reactor - robustness in constraint violations

The second case study of this paper involves a jacketed tubular reactor under steady-state conditions. An irreversible first-order reaction takes place inside the reactor. Two coupled ordinary differential equations are obtained through the mass and energy balances. However, the steady-state scenario is described by an ordinary differential equation in the dimensionless spatial coordinate z denoting the position along the reactor, as the time-dependence is eliminated (Logist et al., 2011).

$$\frac{dx_1}{dz} = \frac{\alpha_{\mathrm{kin}}}{v}(1-x_1)e^{\frac{\gamma x_2}{1+x_2}}, \tag{30}$$

$$\frac{dx_2}{dz} = \frac{\alpha_{\mathrm{kin}}\delta}{v}(1-x_1)e^{\frac{\gamma x_2}{1+x_2}} + \frac{\beta_{\mathrm{kin}}}{v}(u-x_2), \tag{31}$$

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

and with initial conditions:

$$x(0) = (0,0)^\top , \tag{32}$$

and constraints:

$$\frac{T_{\min} - T_{\text{in}}}{T_{\text{in}}} \leq x_2(z) \leq \frac{T_{\max} - T_{\text{in}}}{T_{\text{in}}} , \tag{33}$$

$$\frac{T_{w,\min} - T_{\text{in}}}{T_{\text{in}}} \leq u(z) \leq \frac{T_{w,\max} - T_{\text{in}}}{T_{\text{in}}} . \tag{34}$$

The two states are the dimensionless reactant concentration $x_1 = (C_{\text{in}} - C)/C_{\text{in}}$ and the dimensionless reactor temperature $x_2 = (T - T_{\text{in}})/T_{\text{in}}$. Here, T_{in} and C_{in} are the temperature and the reactant concentration of the feed stream, respectively. The control $u = (T_{\text{w}} - T_{\text{in}})/T_{\text{in}}$ is a dimensionless version of the jacket temperature T_{w}. Both the reactor and jacket temperatures are constrained (Equations (33) and (34)) while the differential equations are solved on the interval $z \in [0,1]$. As OED objective function the D criterion has been chosen. The number of equidistant control intervals is set to 20 and both states are considered to be measurable. The two parameters of interest for the optimal experiment design procedure are $\alpha_{\text{kin}} = 0.058$ and $\beta_{\text{kin}} = 0.2$. The dimensionless version of the reactor jacket temperature u is the only manipulated experimental input. Their assumed parameter variance-covariance matrix is:

$$V = \begin{pmatrix} 0.0174^2 & 0 \\ 0 & 0.06^2 \end{pmatrix} . \tag{35}$$

For the remaining expressions and parameter values, the reader is referred to (Logist et al., 2011).

In a first simulation approach, the parameters are assumed to be normally distributed. Subsequently, the parameters are assumed to be Beta(2,3) distributed with the same mean and variance as the earlier studied normal distribution. Therefore, two stochastic OED approaches are investigated, a first and second order PCE approach based on a normal parametric uncertainty distribution (PCE1 and PCE2) and a first and second order PCE approach based on a Beta(2,3) parametric un-

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

certainty distribution (PCE1 Beta and PCE2 Beta). In particular the following constraints are considered:

$$\mathbb{E}[x_2] + \alpha\sqrt{\mathbb{V}\text{ar}[x_2]} \leq \frac{T_{\max} - T_{\text{in}}}{T_{\text{in}}}, \tag{36}$$

$$\mathbb{E}[x_2] - \alpha\sqrt{\mathbb{V}\text{ar}[x_2]} \geq \frac{T_{\min} - T_{\text{in}}}{T_{\text{in}}}. \tag{37}$$

For the stochastic PCE approaches α is chosen equal to 2 to reduce the number of constraint violations, while for the approximate robust approach a 95% quantile is considered (i.e., $\gamma = 6$).

4.3.1. Normally distributed parameters

When OED is performed the state and control profiles depicted in Figure 4 are obtained. Notice that in the nominal design the maximal temperature never reaches its constraints. The same holds for the approximate robust experiment design. Its corresponding control profile consists of a heating after which a cooling takes place for the remainder of the reactor length. There is a distinct difference between PCE1 and PCE2. The heating profile stops earlier for the PCE1 approach resulting in a reactor temperature which is remarkably lower than the nominal and PCE2 approach. Also note that its upper confidence bound never reaches the state upper bound. In the remainder of the period a cooling takes place however there is a slight risk that the temperature could drop below the lower bound resulting in a reduced cooling effort towards the end of the reactor. For PCE2 the heating is slightly less compared with the nominal case but for its upper confidence region the upper state constraint is active. Subsequently, cooling takes place but similar to the PCE1 approach this is reduced at the end of the simulation interval. Hence, PCE1 a bit more seems conservative, similar to the approximate robust case. However, PCE1 and the approximate robust approach lead to a significantly different result as discussed below.

In order to numerically validate the obtained experiments, 1000 parameter samples are drawn from the assumed Gaussian distribution, subsequently the system is simulated with these parameter values. Given the presence of state bounds, the number of constraint violations is investigated. The simulation results are presented in Table 5. Out of 1000 samples, the nominal D-design results in 15.7% violations.

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

When the experiments obtained by the stochastic approaches are investigated, it is observed that the PCE1 approach for both the 2σ bound results in 14.8%. This result is attributed to the fact that the PCE1 is a coarse approximation of the underlying function. If this does not suffice, the value for α can be increased iteratively or a value can be chosen based on the Cantelli-Chebyshev inequality (Mesbah et al., 2014; Telen et al., 2015). The latter holds no matter what is the underlying distribution of the state bounds. For the PCE2 approaches 5.6% violations are observed. The approximate robust approach results in 3.7% violations in case of normally distributed parameters, which is more robust than most of the stochastic approaches with PCE2.

In Figure 5(a), the valid experiments per designed experiments are illustrated in relation with the sampled parameter values for the PCE2 approach. It clearly depicts which parameter combinations of α_{kin} and β_{kin} yield experiments which violate the state constraints. From the figure it is also clear that there is a set of parameter combinations outside the 95% region which result for the 2σ experiment in a temperature evolution which violates the state constraints.

4.3.2. Beta(2,3) distributed parameters

In the second simulation approach, the parameters are assumed to follow a Beta distribution with the aforementioned mean and standard deviation. Besides mean and variance, a Beta distribution is described by two parameters α_β and β_β which determine the actual shape. For the following simulations $\alpha_\beta = 2$ and $\beta_\beta = 3$ which results in a distribution which is not symmetric with respect to its mean value and has bounded support. It is also apparent in the obtained sampling points to construct the mean and variance approximations as those sampling points are chosen with a higher probability. The same values for α and γ are chosen as for the normally distributed parameters. Note that $\|p - p_{\mathrm{nom}}\|^2_{\Sigma^{-1}}$ is no longer $\chi^2(n_p)$ distributed and that the choice of the quantile γ is in this case not fully correct and only an approximation.

The obtained reactor and jacket temperature profiles are depicted in Figure 4. In contrast with the profiles based on Gaussian distributed parameters from Figure 4,

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

larger confidence intervals are predicted for the expected temperature evolution. This is also evident from the longer extended period in which the confidence interval bounds coincide with the state bounds. In line with the previous simulation, the PCE2 approach reaches a higher reactor temperature and is more constrained by the upper state bound while for PCE1 the lower bound is active. This difference is also apparent in the control action. Note that all Beta distribution based approaches start cooling quicker than the nominal and normally distributed parameters. Note also the difference between PCE1 and PCE2. In the PCE2 case the cooling is not that extreme in order to maintain the upper confidence interval value at the boundary value while PCE1 has some similarity with the observed profiles of the Gaussian case.

Also for this approach, the obtained experiments are validated numerically by sampling 1000 parameter combinations from the assumed Beta distributions. In particular the potential violation of state constraints is once more of interest. The obtained simulation results are depicted in Table 5. For the nominal design violations in 18.1% of the parameter values are observed. For the stochastic approaches, the PCE1 approach is overly robust as even not a single parameter combination resulted in a constraint violation. For the PCE2 approach this is respectively 5.9% and 1.4%. In this second simulation case it is remarkable that the approximate robust approach results in 2.7 % violations, which is less robust than most of the stochastic approaches except the 2 σ experiment with PCE2. Note that the observed robustness of the PCE1 approach in this case study cannot be generalized, i.e., this result is case study specific. From a theoretical point of view, one would expect that lower order PCEs would provide a lower variance (as more positive terms should increase the variance). However, in this case the PCE coefficients are computed with a non-intrusive least-squares apporach which results in an additional source of error. This additional source of error could be on the positive side (i.e., overly robust) or on the negative side (i.e., causing a higher percentage of constraint violations). In this case study the error is on the positive side, such that this overly robust PCE1 result is not problematic.

In Figure 5(b) the valid experiments are depicted for the PCE2 approach in

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

function of the parameter values. Remark here the bounded support for the parameter values. Similar to Figure 5(a) the area where the violations take place is the same.

4.3.3. Computation times

In terms of computation times, similar observations are made as in the Lotka-Volterra case study: the PCE OED approaches require a higher computation time than the approximate robust and nominal OED approaches. The results are summarized in Table 5. Note the factor 2 difference in computation time between the nominal and approximate robust OED approaches, which is most probably due to the nonlinear state constraint. Also remark that PCE1 is a linear approximation but it is significantly more computationally expensive (more than a factor 3) than the approximate robust OED approach. The approximate robust approach in this case study is significantly less computationally expensive than in the Lotka-Volterra case study as the sensitivity states are directly exploited in the evaluation of the robustified constraint. There is no need for additional automatic differentiation to the Jacobian of the Fisher information matrix, as no robustified OED objective is used.

4.3.4. Robustness with respect to information content

Although the focus of this case study is on robustness with respect to constraint violations (feasibility), it is interesting to have a look at the performance of the different OED approaches with respect to robustness with respect to information content. In Table 6 the quartiles and interquartile (IQR) range of the D-criterion are depicted for the Monte Carlo simulations (i.e., the 1000 parameter samples that are drawn from the normal and Beta(2,3) parametric uncertainty distribution, respectively). Note that these quartiles are computed for all samples and that constraint violations are not exclued from these. From this table the following observations can be made. Firstly, the nominal experiment design results in higher information (i.e., higher quartile values) than the stochastic and approximate robust approaches. However, it can be observed that the IQR is smaller than the IQR for the nominal experiment design for the stochastic and approximate robust approaches. This means that in terms of spread of the information content, these approaches perform

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

better than the nominal experiment design. The reduction in information content that can be observed in as well the stochastic as approximate robust approaches is the price to pay for the increased robustness with respect to constraint violations. For this case study, the approximate robust approach performs better in terms of information content than the PCE-based stochastic approaches in case of the Beta(2,3) distributed parametric uncertainty, while for the normally distributed parametric uncertainty, the second order PCE-based approach performs better in terms of information content. Note, however, that for both the normally distributed parameters the number of constraint violations is lower for the approximate robust approach and that in terms of trade-off between information content (optimality) and constraint violations (feasibility) the approximate robust OED approach is recommended for this case study.

4.4. What approach to use for a desired robust experiment design?

In the two presented case studies robustness in information content and robustness in constraint violations have been studied. From the obtained results some guidelines can be formulated with respect to the method that is preferably used when accounting for parametric uncertainty. The guidelines are summarized in Figure 6.

In case that *robustness in information content* is an issue, the PCE based stochastic OED approach is preferred over the approximate robust OED approach. The results clearly indicate that the approximate robust approach is too conservative and results in designs resulting in information content that is four times lower than a nominal OED approach. The stochastic PCE based approach on the other hand results in an improved information content when compared to the nominal and approximate robust approaches. Furthermore, the approximate robust approach requires an additional computational effort in calculating the second order sensitivities for the variance on the objective function. It needs to be noted that with increasing order of PCE the PCE based stochastic OED approaches will have an increased computational cost. However, it is clear that in many cases (as in the presented case studies in this article), a second order PCE is sufficient.

In case that *robustness in constraint violations* is required, the conclusion is not that clear. If computation time is an issue, then the approximate robust approach

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

is preferred over the stochastic PCE based approach as the sensitivity states are directly exploited in the evaluation of the robustified constraint. When considering the percentage of constraint violations, PCE based stochastic OED results in a lower percentage of constraint violations than the approximate robust approach. For the jacketed tubular reactor case study, the reduction in constraint violations is sufficient such that the approximate robust OED approach is preferred in case of robustness in constraint violations.

4.5. Remark

Note that although the PCE approaches clearly perform worse in computation time than the approximate robust approach, the computation time of the stochastic PCE approaches can be reduced by exploiting the sampling-based aspect of the PCE approaches. More specifically, the same dynamic OED system has to be evaluated in the different PCE sampling points. This allows to reformulate the stochastic OED problem with ALADIN (Houska et al., 2016) as a distributed optimization problem in which the different agents consist of the evaluation of the system at different sampling points and the different agents are coupled by the controls (Jiang et al., 2017). Subject of future work will be on an ALADIN reformulation for stochastic optimal control to reduce computational time and construct efficient stochastic optimal control algorithms.

5. Conclusion

The impact of parametric uncertainty on the design of experiments has been studied in this paper. Potential negative effects are an overestimation of the expected information content or experiments that violate operating constraints. In the presented work, a computationally tractable approach based on polynomial chaos expansion has been investigated and compared with the approximate robust optimal experiment design method of Körkel et al. (2004). The presented PCE based approach allows the incorporation of a priori knowledge of the parameter distribution in the uncertainty propagation. In addition, the method allows for a formulation where the expected value and corresponding variance are computed while avoiding a numerical complex integration over the parameter space. The main advantage

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

of the polynomial chaos expansion approach is that more information on the underlying parameter distribution can be incorporated in the optimization problem. The presented PCE based methodology is illustrated with two different case studies with different types of distributions to illustrate the flexibility of the discussed approach and the comparison with the approximate robust optimal experiment design approach of Körkel et al. (2004).

For the Lotka-Volterra case study the presented PCE based methodology is less conservative than the approximate robust methodology and allows to compute information-rich experiments while the variance on the information content is also reduced. The approximate robust approaches lead to a significant loss in information content (almost 80% when compared with nominal experiment designs) and a very small variance on the information content. For the Lotka-Volterra case study, the PCE based stochastic OED formulation is more suitable than the approximate robust OED formulation, since both approaches require a high (and comparable) computation time and the PCE based stochastic OED approach results in a higher information content than the approximate robust approach. In the jacketed tubular reactor case study the emphasis was on reducing constraint violations and generating practically feasible experiments. In case of a normal parametric uncertainty distribution the approximate robust approach resulted in a better reduction of constraint violations than most of the stochastic PCE based methodologies, except the second order PCE approach. However, for the Beta(2,3) approach the stochastic PCE based approaches outperformed the approximate robust approach in terms of constraint violations reduction. However, for the jacketed tubular reactor case study the approximate robust OED approach is more suited since the computation time is much lower than for the PCE based stochastic OED approaches and the number of constraint violations is sufficiently low. The computation time for the approximate robust OED approach is lower than in the Lotka-Volterra case study as no derivative of the Fisher information matrix with respect to the parameters is needed.

A severe limitation of the PCE based stochastic OED formulations is the com-

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

putational cost which increases significantly for an increasing number of states and parameters. However, these formulations exhibit a particular structure originating from the multiple repetitions of the model equations. In future work, the aim is to reformulate the stochastic OED as a distributed optimization problem, consisting of decoupled subsystems. A novel distributed optimization algorithm, ALADIN (Houska et al., 2016),(Jiang et al., 2017), will be used to decouple the large optimization problem and solve the stochastic optimal control problem in a computationally more efficient way. This should allow the application of sampling-based approaches of higher order and to cases with more uncertain parameters and more states. Note however that OED is performed offline so computational time is a hindrance but not a critical issue. From the results obtained for the implemented case studies, it is concluded that the PCE1 approach, due to its first order character, does not always lead to a consistent robustification and is therefore less preferred.

ACKNOWLEDGMENTS

The research was supported by: PFV/10/002 (OPTEC), FWO-G.0930.13 and BelSPO: IAP VII/19 (DYSCO). SB holds a Baekeland PhD grant (03/2016 - 03/2020) from the Agency for Innovation through Science and Technology in Flanders (IWT).

References

Andersson, J., Akesson, J., Diehl, M., 2012. CasADi - a symbolic package for automatic differentiation and optimal control. In: Proceedings of the 6th International Conference on Automatic Differentiation.

Asprey, S., Macchietto, S., 2000. Statistical tools for optimal dynamic model building. Computers and Chemical Engineering 24, 1261 – 1267.

Asprey, S., Macchietto, S., 2002. Designing robust optimal dynamic experiments. Journal of Process Control 12 (4), 545 – 556.

Balsa-Canto, E., Alonso, A., Banga, J., 2008. Computational procedures for optimal experimental design in biological systems. IET Systems Biology 2, 163–172.

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

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.

Bouvin, J., Cajot, S., D'Huys, P.-J., Ampofo-Asiama, J., Anné, J., Van Impe, J., Geeraerd, A., Bernaerts, K., 2015. Multi-objective experimental design for 13c-based metabolic flux analysis. Mathematical Biosciences 268, 22 – 30.

Boyd, S., Vandenberghe, L., 2004. Convex optimization. Cambridge University Press, Cambridge.

Bruwer, M.-J., MacGregor, J. F., 2006. Robust multi-variable identification: Optimal experimental design with constraints. Journal of Process Control 16 (6), 581 – 600.

Cappuyns, A., Bernaerts, K., Smets, I., Ona, O., Prinsen, E., Vanderleyden, J., Van Impe, J., 2007. Optimal fed batch experiment design for estimation of monod kinetics of *Azospirillum brasilense*: From theory to practice. Biotechnology Progress 23 (5), 1074–1081.

Chu, Y., Hahn, J., 2008. Integrating parameter selection with experimental design under uncertainty for nonlinear dynamic systems. AIChE Journal 54 (9), 2310–2320.

Debusschere, B., Najm, H., Pébay, P., Knio, O., Ghanem, R., Maitre, O. L., 2004. Numerical challenges in the use of polynomial chaos representations for stochastic processes. SIAM Journal on Scientific Computing 26, 698–719.

Espie, D., Macchietto, S., 1989. The optimal design of dynamic experiments. AIChE Journal 35, 223–229.

Franceschini, G., Macchietto, S., 2008. Model-based design of experiments for parameter precision: State of the art. Chemical Engineering Science 63, 4846–4872.

Galvanin, F., Barolo, M., Bezzo, F., Macchietto, S., 2010. A backoff strategy for model-based experiment design under parametric uncertainty. AIChE Journal 56, 2088–2102.

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

Häggblom, K. E., 2017. A new optimization-based approach to experiment design for dynamic mimo identification. In: 20th World Congress of The International Federation of Automatic Control, Toulouse, France. pp. 7321 – 7326.

Heine, T., Kawohl, M., King, R., 2008. Derivative-free optimal experimental design. Chemical Engineering Science 63, 4873–4880.

Hjalmarsson, H., 2009. System identification of complex and structured systems. European Journal of Control 15, 275 – 310.

Houska, B., Frasch, J., Diehl, M., 2016. An Augmented Lagrangian Based Algorithm for Distributed Non-Convex Optimization. SIAM Journal on Optimization 26 (2), 1101–1127.

Houska, B., Logist, F., Van Impe, J., Diehl, M., 2012. Robust optimization of nonlinear dynamic systems with application to a jacketed tubular reactor. Journal of Process Control 22, 1152–1160.

Jauberthie, C., Denis-Vidal, L., Coton, P., Joly-Blanchard, G., 2006. An optimal input design procedure. Automatica 42, 881 – 884.

Jiang, Y., Nimmegeers, P., Telen, D., Van Impe, J., Houska, B., 2017. A distributed optimization algorithm for sampling-based stochastic optimal control. In: Submitted to the 20th IFAC World Congress.

Julier, S., Uhlmann, J., 1996. A general method for approximating nonlinear transformations of probability distributions. Tech. rep., Robotics Research Group, Department of Engineering Science, University of Oxford.

Kawohl, M., Heine, T., King, R., 2007. A new approach for robust model predictive control of biological production processes. Chemical Engineering Science 62, 5212 – 5215.

Kiefer, J., Wolfowitz, J., 1959. Optimum designs in regression problems. Annals of Mathematical Statistics 30, 271–294.

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.

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

Liepe, J., Filippi, S., Komorowski, M., Stumpf, M., 2013. Maximizing the information content of experiments in systems biology. PLoS Computational Biology 9, e1002888.

Ljung, L., 1999. System Identification: Theory for the User. Prentice Hall.

Logist, F., Houska, B., Diehl, M., Van Impe, J., 2011. Robust multi-objective optimal control of uncertain (bio)chemical processes. Chemical Engineering Science 66 (20), 4670 – 4682.

Mesbah, A., Streif, S., 2015. A probabilistic approach to robust optimal experiment design with chance constraints. In: International Symposium on Advanced Control of Chemical Processes (ADCHEM), 2015. pp. 100–105.

Mesbah, A., Streif, S., Findeisen, R., Braatz, R., 2014. Stochastic nonlinear model predictive control with probabilistic constraints. In: Proceedings of the American Control Conference (ACC). pp. 2413–2419.

Mitra, K., 2009. Multiobjective optimization of an industrial grinding operation under uncertainty. Chemical Engineering Science 64 (23), 5043–5056.

Nagy, Z., Braatz, R., 2004. Open-loop and closed-loop robust optimal control of batch processes using distributional and worst-case analysis. Journal of Process Control 14, 411–422.

Nagy, Z., Braatz, R., 2007. Distributional uncertainty analysis using power series and polynomial chaos expansions. Journal of Process Control 17, 229–240.

Nimmegeers, P., Telen, D., Logist, F., Van Impe, J., 2016. Dynamic optimization of biological networks under parametric uncertainty. BMC Systems Biology 10 (1), 86.

Nimmegeers, P., Telen, D., Van Impe, J., 2017. A sampling-based stochastic optimal experiment design formulation with application to the williams-otto reactor. In: 20th World Congress of The International Federation of Automatic Control, Toulouse, France. pp. 9046–9051.

Paulson, J. A., Buehler, E., Mesbah, A., 2017. Arbitrary polynomial chaos for uncertainty propagation of correlated random variables in dynamic systems. In: 20th

World Congress of The International Federation of Automatic Control, Toulouse, France. pp. 3607–3612.

Pronzato, L., Walter, E., 1985. Robust experiment design via stochastic approximation. Mathematical Biosciences 75, 103–120.

Pronzato, L., Walter, E., 1988. Robust experiment design via maximin optimization. Mathematical Biosciences 89, 161 – 176.

Recker, S., Kühl, P., Diehl, M., Bock, H., 2012. Sigmapoint approach for robust optimization of nonlinear dynamic systems. In: Proceeding of SIMULTECH 2012. pp. 199–207.

Rojas, C., Goodwin, J. W. G., Feuer, A., 2007. Robust optimal experiment design for system identification. Automatica 43, 993–1008.

Sager, S., 2013. Sampling decisions in optimum experimental design in the light of Pontryagin's maximum principle. SIAM Journal on Control and Optimization 51, 3181–3207.

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.

Schenkendorf, R., Kremling, A., Mangold, M., 2009. Optimal experimental design with the sigma point method. IET Systems Biology 3, 10–23.

Srinivasan, B., Bonvin, D., Visser, E., Palanki, S., 2003. Dynamic optimization of batch processes II. Role of measurements in handling uncertainty. Computers and Chemical Engineering 27, 27–44.

Telen, D., Houska, B., Logist, F., Diehl, M., Van Impe, J., 2013a. Guaranteed robust optimal experiment design for nonlinear dynamic systems. In: Proceedings of the 12th European Control Conference. Zurich, Switserland, pp. 2939–2944.

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

Telen, D., Houska, B., Logist, F., Vanderlinden, E., Diehl, M., Van Impe, J., 2013b. Optimal experiment design under process noise using Riccati differential equations. Journal of Process Control 23, 613–629.

Telen, D., Logist, F., Van Derlinden, E., Van Impe, J., 2012a. A probabilistic approach to robust optimal experiment design with chance constraints. In: 20th Mediterranean Conference on Control & Automation (MED), 2012. pp. 157–162.

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

Telen, D., Vallerio, M., Cabianca, L., Houska, B., Van Impe, J., Logist, F., 2015. Approximate robust optimization of nonlinear systems under parametric uncertainty and process noise. Journal of Process Control 33, 140 – 154.

Telen, D., Vercammen, D., Logist, F., Van Impe, J., 2014. Robustifying optimal experiment design for nonlinear, dynamic (bio)chemical systems. Computers and Chemical Engineering 71, 415–425.

Vassiliadis, V., Balsa-Canto, E., Banga, J., 1999. Second-order sensitivities of general dynamic systems with application to optimal control problems. Chemical Engineering Science 54, 3851–3860.

Wächter, A., Biegler, L., 2006. On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. Mathematical Programming 106 (1), 25–57.

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

Welsh, J. S., Rojas, C. R., 2009. A scenario based approach to robust experiment design. IFAC Proceedings Volumes 42 (10), 186 – 191, 15th IFAC Symposium on System Identification.

Wendt, M., Li, P., Wozny, G., 2002. Nonlinear chance-constrained process optimization under uncertainty. Industrial and Engineering Chemistry Research 41, 3621–3629.

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

Wiener, N., 1938. The homogeneous chaos. American Journal of Mathematics 60 (4), 897 – 936.

Xiu, D., Karniadakis, G., 2002. The wiener-askey polynomial chaos for stochastic differential equations. SIAM Journal of Scientific Computation 24, 619–644.

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

Tables

Table 1: Objective and constraint formulations for the nominal, approximate robust and PCE based stochastic optimal experiment design approaches.

	Nominal OED	Approximate robust OED	PCE based stochastic OED
Objective	$\Phi(F(t_{\mathrm{f}}))$	$\Phi(F(t_{\mathrm{f}})) + \sqrt{\gamma}\left\Vert \frac{d}{dp}\Phi(F(t_{\mathrm{f}}))\right\Vert_{\Sigma}$	$a^{d}_{\Phi,0} + \alpha_{\Phi}\sqrt{\sum_{j=1}^{L-1}\left(a^{d}_{\Phi,\mathrm{j}}\right)^{2}\mathbb{E}\left[\Psi^{2}_{\mathrm{j}}(p)\right]}$
Constraint	$c_i(x(t),u(t),t)$	$c_i(x(t),u(t),t) + \sqrt{\gamma}\sqrt{\left(\frac{dx}{dp}\right)^{\top}\left(\frac{dc_i}{dx}\right)^{\top}\Sigma\frac{dc_i}{dx}\frac{dx}{dp}}$	$a^{d}_{c_i,0} + \alpha_{c_i}\sqrt{\sum_{j=1}^{L-1}\left(a^{d}_{c_i,\mathrm{j}}\right)^{2}\mathbb{E}\left[\Psi^{2}_{\mathrm{j}}(p)\right]}$

Table 2: E-criterion values obtained from simulating the model with the current best guess of the parameter values and the controls from the nominal and expected E designs for the approximate robust, PCE1 and PCE2 approaches ($\alpha = 1.65$).

	Nominal	Robust	PCE1	PCE2	PCE2 Uniform
λ_{min}	45.55	8.90	30.97	44.31	39.62
$\lambda_{\mathrm{min}}/\lambda_{\mathrm{min,nom}}$	1	0.195	0.680	0.973	0.870

Table 3: Quartiles E-criterion value for Monte Carlo simulations with 1000 realizations from normally, uniformly distributed parameters p_1 and p_2 with mean 1 and standard deviation 0.1 for $\alpha = 1.65$.

	Nominal	Robust	PCE1	PCE2	PCE2 uniform
Q1 Normal	31.507	8.662	23.958	33.327	-
Q2 Normal	39.336	9.877	29.401	40.964	-
Q3 Normal	49.579	12.285	37.094	51.522	-
Q1 Uniform	30.728	8.733	23.716	32.768	31.366
Q2 Uniform	38.728	10.342	29.550	41.094	39.695
Q3 Uniform	51.139	12.385	38.890	53.336	52.177

Table 4: Computation times for the Lotka-Volterra case study required for nominal OED, approximate robust OED stochastic EV OED and stochastic robustified OED in seconds.

Nominal	Robust	PCE1 EV	PCE2 EV	PCE2 uniform EV
91.06	553.97	799.93	3869.12	3445.1
		PCE1 1.65	PCE2 1.65	PCE2 uniform 1.65
		4051.22	7355.3	4447.42

Table 5: Number and percentage of constraint violations, computation times and number of states for the different experiment designs for the jacketed tubular reactor case study.

	Normally distributed parameters			
	Nominal	Robust	PCE1	PCE2
$n_{\mathrm{violations}}$	157	37	148	56
% violations	15.7	3.7	14.8	5.6
CPU time	13.93	24.4	87.13	428.21
	Beta(2,3) distributed parameters			
	Nominal	Robust	PCE1 Beta	PCE2 Beta
$n_{\mathrm{violations}}$	181	25	0	59
% violations	18.1	2.5	0	5.9
CPU time	13.93	24.4	115.67	447.92

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

Table 6: Quartiles (Q1, Q2 and Q3) and interquartile range (IQR=Q3-Q1) D-criterion value for Monte Carlo simulations with 1000 realizations from normally, Beta(2,3) distributed parameters.

	Nominal	Robust	PCE1 2σ	PCE2 2σ
Q1 Normal	0.281	0.158	0.153	0.184
Q2 Normal	0.473	0.293	0.299	0.346
Q3 Normal	0.806	0.497	0.548	0.619
IQR Normal	0.525	0.339	0.395	0.434
Q1 Beta(2,3)	0.263	0.145	0.037	0.142
Q2 Beta(2,3)	0.470	0.283	0.085	0.264
Q3 Beta(2,3)	0.820	0.570	0.220	0.476
IQR Beta(2,3)	0.557	0.426	0.183	0.333

Figures

Journal homepage: https://www.journals.elsevier.com/chemical-engineering-science
Original file available at: https://doi.org/10.1016/j.ces.2020.115651

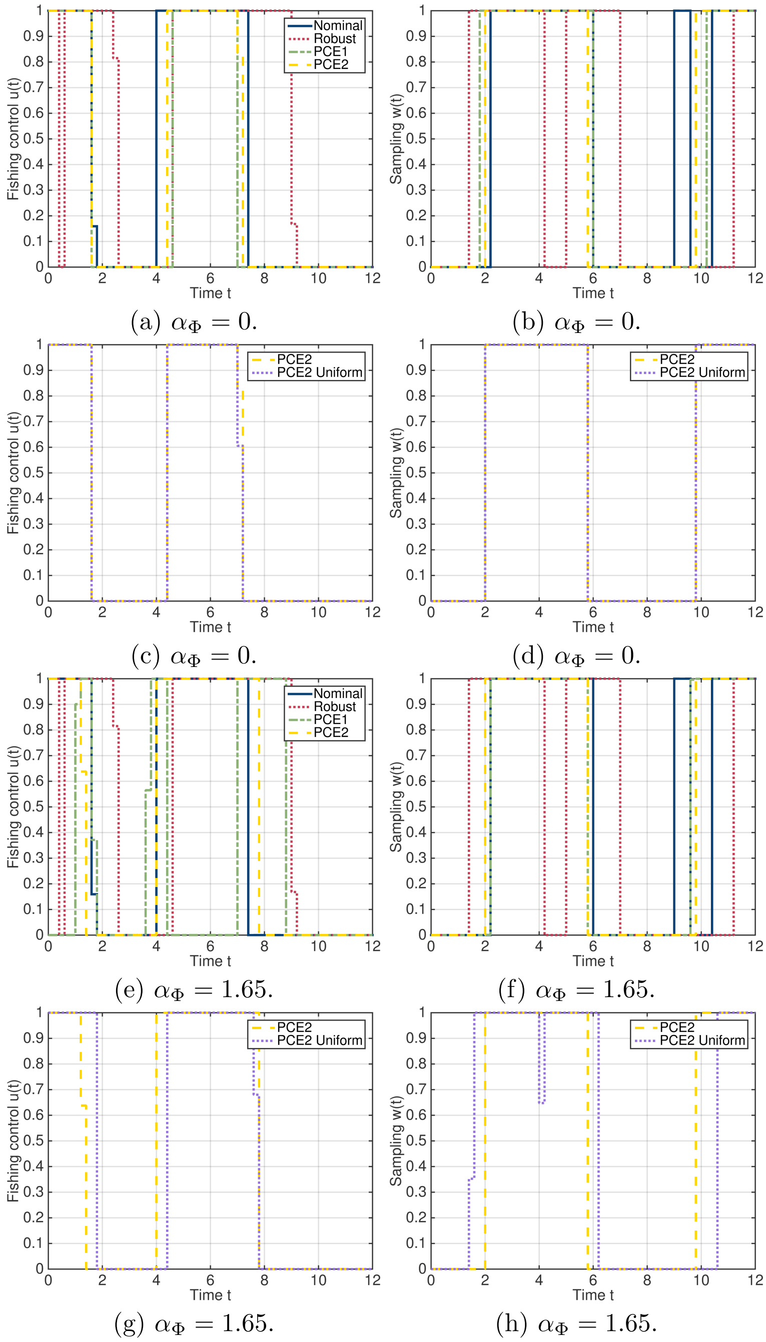
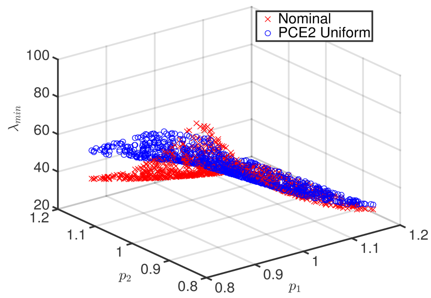
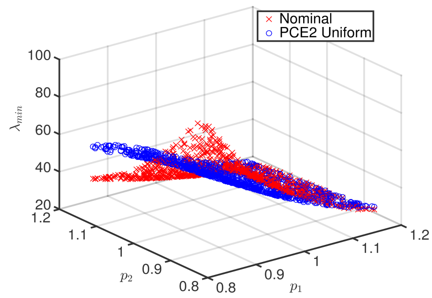


Figure 1: Fishing control $u(t)$ and sampling action $w(t)$ profiles for the E-optimal, approximate robust (Robust), and PCE-based stochastic (PCE1, PCE2 and PCE2 Uniform) experiment designs.

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`



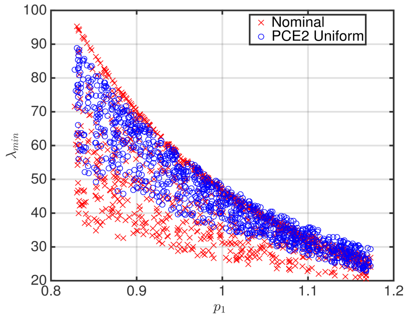
(a) $\alpha = 0$.



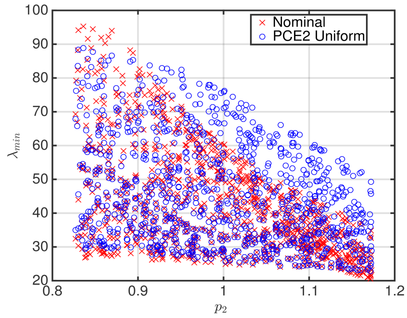
(b) $\alpha = 1.65$

Figure 2: Depiction of the two minimum eigenvalue surfaces for the nominal case and the PCE2 approach for a uniform distribution.

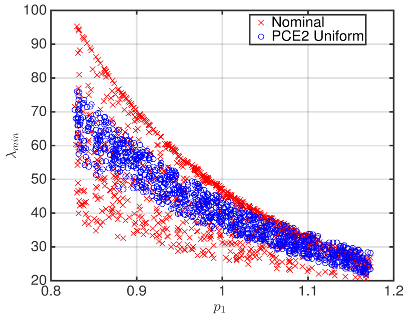
Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`



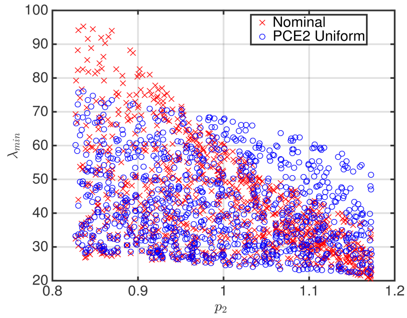
(a) Minimum eigenvalue surface in function of parameter $p1$ for $\alpha = 0$.



(b) Minimum eigenvalue surface in function of parameter $p2$ for $\alpha = 0$.



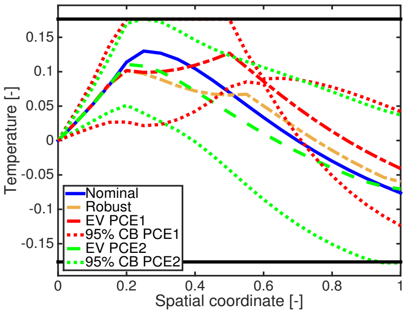
(c) Minimum eigenvalue surface in function of parameter $p1$ for $\alpha = 1.65$.



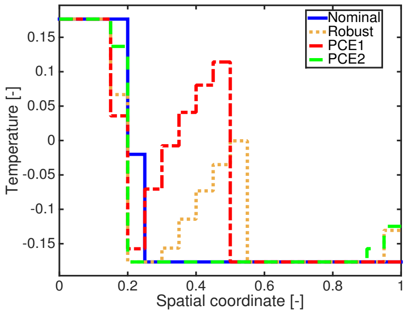
(d) Minimum eigenvalue surface in function of parameter $p2$ for $\alpha = 1.65$.

Figure 3: The 2 two dimensional projections of Figure 2.

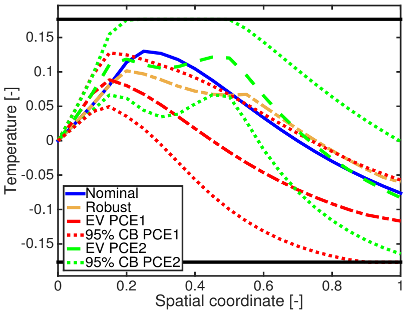
Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`



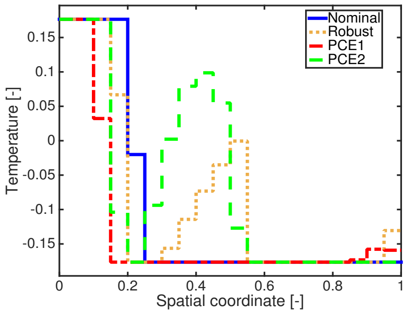
(a) Reactor temperature for normally distributed parameters.



(b) Cooling jacket temperature for normally distributed parameters.

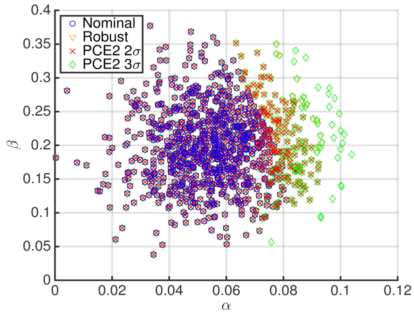


(c) Reactor temperature for Beta(2,3) distributed parameters.

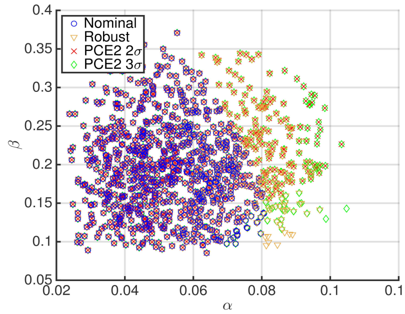


(d) Cooling jacket temperature for Beta(2,3) distributed parameters.

Figure 4: Simulated reactor temperature evolution with 95% confidence bound and control actions of the D-design and two stochastic OED designs for normally distributed parameters (a,b) and Beta distributed parameters (c,d).



(a) Normally distributed parameters.



(b) Beta(2,3) distributed parameters.

Figure 5: Depiction of valid experiments out of 1000 parameter samples for each of the different designed experiments based on the PCE2 approach of the jacketed tubular reactor.

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

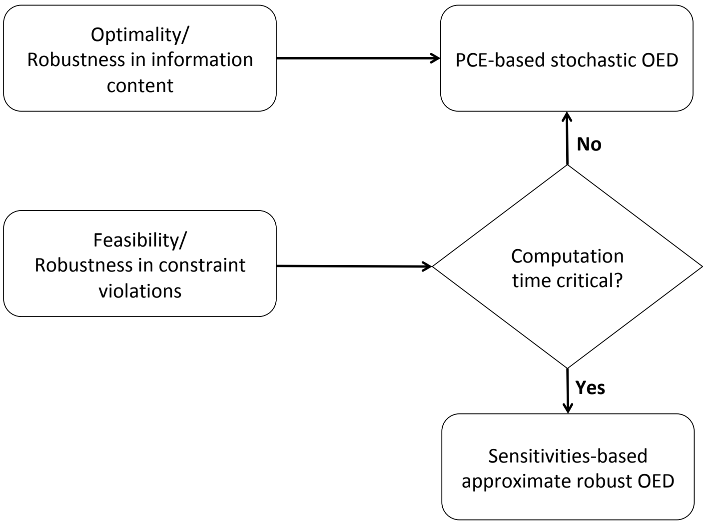


Figure 6: Decision tree to select OED approach based on desired robustness, i.e., with respect to information content (optimality) or with respect to constraint violations (feasibility).

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

Appendix A. Derivation of the solution for the linearized inner problem

Consider the inner maximization problem which is approximated with a first order Taylor expansion in the variable p and reformulated as:

$$\max_{p} \Phi(F(t_{\mathrm{f}})) + \frac{d}{dp}\Phi(F(t_{\mathrm{f}}))(p - p_{\mathrm{nom}}) \tag{A.1}$$

$$\text{subject to: } \|p - p_{\mathrm{nom}}\|^2_{\Sigma^{-1}} \leq \gamma \tag{A.2}$$

It is clear from this formulation that the objective function of this maximization problem is linear in the parameters p. Since $\Phi(F(t_{\mathrm{f}}))$ is evaluated in the nominal parameter values p_{nom}, the only relevant term for the maximization is $\frac{d}{dp}\Phi(F(t_{\mathrm{f}}))(p - p_{\mathrm{nom}})$. Hence the maximization problem can be simplified to:

$$\max_{p} \frac{d}{dp}\Phi(F(t_{\mathrm{f}}))(p - p_{\mathrm{nom}}) \tag{A.3}$$

$$\text{subject to: } \|p - p_{\mathrm{nom}}\|^2_{\Sigma^{-1}} \leq \gamma \tag{A.4}$$

Consider the Lagrangian $\mathbb{L}$ of Equation (A.3):

$$\mathbb{L} = \Phi(F(t_{\mathrm{f}})) + \frac{d}{dp}\Phi(F(t_{\mathrm{f}}))(p - p_{\mathrm{nom}}) + \lambda(\gamma - (p - p_{\mathrm{nom}})^{\top}\Sigma^{-1}(p - p_{\mathrm{nom}})), \quad \lambda \geq 0 \tag{A.5}$$

From differentiation with respect to the optimization variable p and the necessary optimality condition $\frac{d\mathbb{L}}{dp} = 0$:

$$\frac{d}{dp}\Phi(F(t_{\mathrm{f}})) + 2\lambda(\Sigma^{-1}p_{\mathrm{nom}} - \Sigma^{-1}p) = 0 \tag{A.6}$$

If $\lambda = 0$, $\frac{d}{dp}\Phi(F(t_{\mathrm{f}}))(p - p_{\mathrm{nom}})$ would be independent of p. This is a contradiction and would make the optimization problem irrelevant. Hence, $\lambda > 0$. Consequently, Equation (A.6) can be reformulated to an expression for p:

$$p = p_{\mathrm{nom}} + \frac{1}{2\lambda}\Sigma\frac{d}{dp}\Phi(F(t_{\mathrm{f}})) \tag{A.7}$$

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`

Since Equation (A.1) is a linear, quadratic constrained problem in p and $\lambda > 0$, the optimal solution is at the boundary (Boyd and Vandenberghe, 2004), i.e.,

$$(p - p_{\text{nom}})^\top \Sigma^{-1} (p - p_{\text{nom}}) = \gamma \tag{A.8}$$

Substituting Equation (A.7) in Equation (A.8) results in the following expression:

$$\left(\frac{1}{2\lambda}\Sigma \frac{d}{dp}\Phi(F(t_{\text{f}}))\right)^\top \Sigma^{-1} \frac{1}{2\lambda}\Sigma \frac{d}{dp}\Phi(F(t_{\text{f}})) = \gamma \tag{A.9}$$

From Equation (A.9) and $\lambda > 0$, an optimal solution for λ is determined:

$$\lambda = \frac{1}{2\sqrt{\gamma}} \sqrt{\left(\frac{d}{dp}\Phi(F(t_{\text{f}}))\right)^\top \Sigma \frac{d}{dp}\Phi(F(t_{\text{f}}))} \tag{A.10}$$

The optimal p is given by:

$$p = p_{\text{nom}} + \frac{\sqrt{\gamma}}{\sqrt{\left(\frac{d}{dp}\Phi(F(t_{\text{f}}))\right)^\top \Sigma \frac{d}{dp}\Phi(F(t_{\text{f}}))}} \Sigma \frac{d}{dp}\Phi(F(t_{\text{f}})) \tag{A.11}$$

Evaluation of the objective function of the inner maximization problem in the optimal solution p leads to:

$$J_{\text{inner}} = \Phi(F(t_{\text{f}})) + \frac{d}{dp}\Phi(F(t_{\text{f}})) \frac{\sqrt{\gamma}\Sigma \frac{d}{dp}\Phi(F(t_{\text{f}}))}{\sqrt{\left(\frac{d}{dp}\Phi(F(t_{\text{f}}))\right)^\top \Sigma \frac{d}{dp}\Phi(F(t_{\text{f}}))}} \tag{A.12}$$

$$= \Phi(F(t_{\text{f}})) + \sqrt{\gamma \frac{d}{dp}\Phi(F(t_{\text{f}}))\Sigma \frac{d}{dp}\Phi(F(t_{\text{f}}))} \tag{A.13}$$

$$= \Phi(F(t_{\text{f}})) + \sqrt{\gamma} \left\| \frac{d}{dp}\Phi(F(t_{\text{f}})) \right\|_{\Sigma} \tag{A.14}$$

Journal homepage: `https://www.journals.elsevier.com/chemical-engineering-science`
Original file available at: `https://doi.org/10.1016/j.ces.2020.115651`